TAF10 interacts with GATA1 transcription factor and controls mouse erythropoiesis

Petros Papadopoulos[a,f]#, Laura Gutiérrez[a,d], Jeroen Demmers[b], Elisabeth Scheer[c], Farzin Pourfarzad[a,d], Dimitris N. Papageorgiou[e], Elena Karkoulia[e], John Strouboulis[e], Harmen J.G. van de Werken[a], Reinier van der Linden[a], Peter Vandenberghe[f], Dick H.W. Dekkers[b], Sjaak Philipsen[a], Frank Grosveld[a] and Làszlò Tora[c]#

Department of Cell Biology, Erasmus MC, Rotterdam, The Netherlands[a], Proteomics Center, Erasmus MC, Rotterdam, The Netherlands[b], Cellular Signaling and Nuclear Dynamics Program, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), UMR 7104 CNRS, INSERM U964, Université de Strasbourg, Illkirch, France[c], Department of Blood Cell Research, Sanquin Research and Landsteiner Laboratory, AMC, Amsterdam, The Netherlands[d], Division of Molecular Oncology, B.S.R.C "Alexander Fleming", Athens, Greece[e], Center for Human Genetics, KULeuven, Leuven, Belgium[f]

Running Title: TAF10 controls mouse erythropoiesis

#Correspondence should be addressed to:
Làszlò Tora, email: laszlo@igbmc.fr; Phone (LT): +33 388653444, Fax: +33-388653201
Petros Papadopoulos, email: petros.papadopoulos@med.kuleuven.be; Phone (PP): +32 16347992, Fax: +32-16346082

Key words: transcription, TAF10, GATA1, erythropoiesis, TFIID, SAGA

Abstract word counts:	194
Materials and Methods word counts:	955
Text word counts:	6033

Abstract

The ordered assembly of a functional pre-initiation complex (PIC), composed of general transcription factors (GTFs), is a prerequisite for the transcription of protein-coding genes by RNA polymerase II. TFIID, comprised of the TATA binding protein (TBP) and 13 TBP-associated factors (TAFs), is the GTF that is thought to recognize the promoter sequences allowing site-specific PIC assembly. Transcriptional cofactors, such as SAGA, are also necessary for tightly regulated transcription initiation. The contribution of the two TAF10-containing complexes (TFIID, SAGA) to erythropoiesis remains elusive. By ablating TAF10 specifically in erythroid cells *in vivo* we observed a differentiation block accompanied by deregulated GATA1 target genes, including *Gata1* itself, suggesting functional crosstalk between GATA1 and TAF10. Additionally, we analyzed the composition of TFIID and SAGA complexes by mass spectrometry in mouse and human cells and found that their global integrity is maintained, with minor changes, during erythroid differentiation and development. In agreement with our functional data, we show that TAF10 interacts directly with GATA1 and that TAF10 is enriched on the GATA1 locus in human fetal erythroid cells. Thus, our findings demonstrate a crosstalk between canonical TFIID and SAGA complexes and cell-specific transcription activators during development and differentiation.

Molecular and Cellular Biology

Introduction

Initiation of RNA polymerase II (RNA pol II) transcription in eukaryotes is a process involving stepwise recruitment and assembly of the preinitiation complex (PIC) at the core promoter of a transcriptional unit. Transcription factor TFIID, comprised of the TATA binding protein (TBP) and a series of TBP-associated factors (TAFs), is the general transcription factor (GTF) that by recognizing the promoter sequences and surrounding chromatin marks, allows the site-specific assembly of the PIC(1) (and refs therein). Binding of the TFIID complex is aided by TFIIA and followed by the remainder of the general transcription machinery, including TFIIB, RNA pol II/TFIIF, TFIIE and TFIIH complexes. Additional cofactors, including the Mediator complex, histone modifiers and chromatin remodelers, facilitate the communication between gene-specific transcription factors and the general transcription machinery.

The TFIID complex does not only bind to TATA box–containing promoters, but also to TATA-less promoters and this led to the idea that TAFs could provide TFIID with additional functional features(2, 3). Indeed, nine out of thirteen TAFs contain a histone fold domain (HFD)(4) favoring the formation of TAF heterodimers. For instance, the TAF6-9 heterodimer has been found to bind promoter elements downstream of the TATA-box(5-7) and is a direct target of transcriptional activators(8, 9). Moreover, it has been shown that TAF knockouts (KO) and *in vitro* TAF knock-down experiments result in both down- and up-regulated gene expression of subsets of genes(10, 11). All these results together suggest that TFIID is a highly flexible regulator of transcription, functioning both in gene activation and in repression.

Additionally, co-activator complexes with histone acetyltransferase (HAT) activity responsible for gene activation-associated interactions, including the ATAC (Ada-Two-A-Containing) and SAGA (Spt-Ada-Gcn5-acetyltransferase) complexes, appear to have

distinct functional roles by targeting either promoters or enhancers or both (12) (and refs therein).

TAF10 is a subunit of both the TFIID and the SAGA co-activator HAT complexes(13). The role of TAF10 is indispensable for early embryonic transcription and mouse development as knockout (KO) embryos die early in gestation between E3.5 and E5.5, around the stage when the supply of maternal protein becomes insufficient(14). However, when analyzing TFIID stability and transcription it was noted that not all cells and tissues were equally affected by the loss of TAF10. For example, ablation of TAF10 in keratinocytes impaired skin barrier formation and deregulated a subset of related genes when inactivated during the fetal stage, but resulted in no detectable effect in adult keratinocytes(15). Moreover, studies in which TAF10 was conditionally ablated in fetal or adult liver demonstrated the essential role of TAF10 in liver development revealing the necessity of TAF10 for TFIID stability to repress specific genes in the liver in postnatal life(10). These findings together confirm that TAF10, probably as a subunit of TFIID and/or SAGA, is essential during mouse development and suggest that TAF10 plays an important role during embryonic development and homeostasis in a tissue-dependent manner. Understanding the interplay of TAF10-containing TFIID and SAGA complexes with developmentally important and tissue-specific transcription factors is crucial to obtain a more comprehensive view of cell differentiation throughout development.

Erythropoiesis is the process by which red cells are formed(16). There are two waves of erythropoiesis in mammals, primitive and definitive. Definitive erythropoiesis starts in the fetal liver, and moves later during gestation to the spleen and bone marrow, which in mice remain the sites of erythropoiesis during adulthood. The fetal and adult stages of definitive erythropoiesis differ at the transcriptional level, exemplified in humans by the type of beta-hemoglobin chain expressed. Many tissue-specific transcription factors have been studied in order to provide mechanistic clues to this process of developmental-stage specific

Molecular and Cellular Biology

hemoglobin expression(17). GATA1 is one of them, it is expressed in lineage-committed cells (erythroid, megakaryocytic, eosinophilic, mast and dendritic cells) and plays an important role in the regulation of differentiation and survival of these lineages(18, 19). Embryos lacking *Gata1* die at approximately embryonic day 11.5 due to the maturation arrest of primitive erythroid cells(20), while conditional knockout of *Gata1* in adults leads to red cell aplasia and severe thrombocythopenia(21).However, the composition of general transcription complexes, such as TFIID and SAGA, and the role of TAFs during developmental erythropoiesis have not been investigated yet.

Alternative TFIID and other GTF complexes have been implicated in providing alternative pathways leading to gene regulation during differentiation(22, 23). To gain insight into the role of GTF in mouse erythropoiesis we have carried out specific inactivation of TAF10, a cornerstone subunit of the TFIID and SAGA complexes(10), in the erythroid compartment by crossing *TAF10lox*(14) with *EpoR-Cre* mice(24). We found that TAF10 ablation results in a block of erythropoiesis leading to severe anemia, which is lethal at E13.5. Several GATA1 target genes, including *Gata1* itself, were deregulated when TAF10 was ablated. We have also analyzed the composition and stoichiometry of the TAF10-containing transcription complexes, TFIID and SAGA, by mass spectrometry in proliferating erythroid precursors (proerythroblasts) and synchronously differentiated erythroid cells of mouse and human origin. Interestingly, we found that TAF10 interacts physically with the master regulator of erythroid differentiation, the GATA1 transcription factor, and we observed enrichment of TAF10 binding to the *GATA1* locus in human erythroid cells. Collectively, these data suggest that the interaction of TAF10 with GATA1 is important to facilitate the recruitment of TFIID and/or SAGA to GATA1-responsive promoters and that the autologous control of GATA1 expression(25) requires the presence of TAF10 in these complexes.

Molecular and Cellular Biology

Materials and Methods

Mice

The *TAF10lox/KO*(14) mice and *EpoR-Cre*(24) mice have been described previously. *TAF10lox/KO:EpoR-Cre*$^{+/-}$ (TAF10KO^{cEry}) and single genetically modified mice were maintained in a C57BL/6 background. All experiments described in this article have been approved and conducted according to the animal welfare committee guidelines of Erasmus MC Rotterdam.

Human material

Samples of fetal liver and buffy coats from peripheral blood were provided by the clinic and the Sanquin Blood bank in The Netherlands, in compliance with the Erasmus MC Ethical Guidelines Committee.

Flow cytometry

Fetal livers were dissected from embryos at E11.5, E12.5 and E13.5, and resuspended in phosphate buffer saline (PBS) supplemented with 1% bovine serum albumin (BSA) to single cell suspensions. Cells were stained on ice for 20 min and washed twice with PBS/1%BSA before acquisition on a FacsARIAII instrument (Becton Dickinson, BD). Antibodies used for staining include, TER119-PerCP-Cy5.5 (BD), KIT (CD117)-PECy7 (BD), CD71-PE (BD). Since the *EpoR-Cre* allele includes a GFP reporter(24), EpoR-driven Cre expression was followed by GFP in the FITC channel. Dead cells were excluded by Hoechst staining (Invitrogen) and analysis was performed with the FlowJo software (Tree Star, Inc).

Extract preparation and Immunoprecipitation (IP) Assays

Nuclear extracts were prepared as described(10, 26). TAF10 immunoprecipitations (IPs) were performed as described for mouse(10) and human cells(27). Mouse monoclonal antibodies, 23TA 1H8(28) and 6TA 2B11(14), were used for human(29) and mouse TAF10 IPs, respectively, and a mix of 6TA 2B11 and 6TA 4G2 (27) (1/1000) against TAF10 protein

for assessment of IP efficiency by Western blotting. Mock IP was performed using a GST antibody (Santa Cruz, sc-80004) according to the manufacturer's recommendations.

Mass spectrometry was performed as described before(30) on a LTQ-Orbitrap mass spectrometer (Thermo).

Purification of proteins

GST-GATA1 and GST proteins (pGEX expression vectors) were expressed in BL21 *E. coli* upon IPTG (0.4 mM) induction for 3 h at 30 °C. Bacteria were lysed (50 mM Tris-HCl pH 8.0, 400 mM NaCl, 5% Glycerol, 1mM Glutathione, 2,5 mM PMSF, 50 µg/ml DNAse and 1 mM $MgCl_2$, complete protease inhibitor cocktail (Roche)) and passed through the pressure plunge three to four times. After centrifugation at 20.000 rpm for 20 min at 4 °C the extract was incubated with pre-equilibrated (wash buffer without glutathione) glutathione beads (Glutathione Sepharose 4B, GE Healthcare) for 1 h at room temperature (RT). After three washes (50 mM Tris-HCl pH 8.0, 1 M NaCl, 5% Glycerol, 1 mM Glutathione, complete protease inhibitor cocktail) the proteins were eluted (50 mM Glutathione, 50 mM Tris-HCl pH 8.0, 150 mM NaCl, complete protease inhibitor cocktail).

Chromatin immunoprecipitation (ChIP)

Chromatin preparation from human erythroid progenitor cultures (HEPs) was previously described(10). Fetal liver and adult HEPs were cultured(31) and used for ChIP reactions that were performed as described(32) with the 23TA 1H8 antibody clone against the human TAF10 protein, the GATA1 antibody (ab11852) and a CD71 antibody (347510, BD Biosciences) as a negative control. qPCR was performed on the input and immunoprecipitated samples using primers for the HS2 (palindromic GATA1 binding site) or a more proximal GATA1 binding site at the GATA1 promoter. The relative fold-enrichment was calculated using the delta delta Ct method(33) , setting the relative fold-enrichment of CD71 background binding to 1. Primers used are Hs GATA1 palindromic binding site Fw 5'-AGACTTATCTGCTGCCCCAG-3', Hs GATA1 palindromic binding site Rev 5'-

CCAGGCTAAGCCTGCAGGC-3' or Rev 5'-TAGAGCCTGTGGGATACCTTG-3'; Hs GATA1 binding site at -3kb Fw 5'- GGGATGAGGGAATAGTGGTG-3', Hs GATA1 binding site at -3kb Rev 5'- GCTCTTTGTCTCTGTGTCTCTGTC-3'.

Gene expression (RNA-seq and qRT-PCR)

The RNA-seq library was generated according to the Illumina protocol using the TruSeq RNA Sample Prep Kit v2 kit. 500ng of total RNA was initially extracted with Trizol Reagent (Invitrogen) from mouse fetal livers of TAF10KO and WT embryos at E12.5. Quality of RNA was checked with a Nanodrop analyzer and on an Agilent Technologies Bioanalyzer 2100. Sequencing was performed on an Illumina HiSeq 2000 instrument. The reads were aligned with Tophat version 2.0.8 [http://genomebiology.com/2013/14/4/R36/abstract] on the UCSC mm10 reference genome using the Ensembl version 75 gene models(34).

qRT-PCR

cDNA synthesis was performed from 1µg of RNA with the DyNamo cDNA synthesis kit (F-470, Finnzymes) and diluted to 10ng/µl concentration from which 10ng was used per reaction at 20µl total volume. Conditions of the qRT-PCR and primer sequences are already described(10) and reactions were performed in Applied Biosystems Thermal Real Time PCR instruments (ABI 7900).

RNA-Seq analysis

Raw counts were measured with htseq-count version 0.6.0. using -m union -s no -a 20 as settings(35). The counted data were normalized by the size factor of the libraries and subsequently the Fragments Per Kilobase of exon per Million fragments mapped (FPKM) were calculated. The differentially expressed genes were called using a generalized linear negative binomial model that controlled for the effect of the RNA sample preparation date. The calculations were performed by the DESeq2 R package(36). The False Discovery Rate (FDR) was calculated with Benjamini Hochberg method(37) and the threshold value was set to 0.01. Gene Ontology (GO) gene-enrichment analyses were carried out with the GO-stat

package(38) using a conditional hypergeometrical test for over-represented Ensembl Gene IDs using a threshold *p*-value of 1 x 10^{-6}. The Principal Component Analysis (PCA) was carried out using the Pearson's correlation matrix after blind variance stabilizing transformation of the normalized counts(36).The RNA-seq data was joined with GATA1 ChIP-seq data(39) and plotted in a Venn diagram. The statistical package R was used for calculations and plotting of the data (R Core Team (2014). R: A Language and Environment for Statistical Computing. *Vienna, Austria: R Foundation for statistical computing*; http://www.r-project.org.)

Culture of erythroid cells

Erythroid cell culture conditions have been described for mouse(40) and human erythroid progenitors(31). Cells were kept under proliferation or differentiation conditions, and they were followed by cell density and size monitoring with a CASY instrument (Innovatis, Roche Diagnostics GmbH).

Molecular and Cellular Biology

Results

Erythroid-specific knockout of TAF10 shows its requirement for erythropoiesis

To better understand the role of TFIID and SAGA in erythroid differentiation we generated mice in which *TAF10* was specifically ablated in erythroid cells. *TAF10+/-* mice are normal(14), therefore we generated mice bearing a *TAF10* knockout allele and a *TAF10* LoxP-flanked allele in order to reduce the recombination requirements to one allele when generating conditional knockout mice. In order to study TAF10 loss-of-function specifically in erythroid cells we generated *TAF10KO/Lox* animals bearing the erythroid-specific *EpoR-Cre* allele(24), and we refer to them as *TAF10KO*^{cEry}. The expression of EpoR starts at E8.0 in yolk sac erythroid progenitors (41), increasing during definitive erythropoiesis in BFUe and reaching maximal levels in CFUe progenitors.

Definitive erythropoiesis starts at E10.5 in the fetal liver but definitive erythroid cells are detected in the circulation after E11.5, and only at E13.5 reach an approximate 1:1 ratio with still circulating primitive erythroid cells derived from the yolk sac (42, 43). In order to analyze this developmental transition period, we performed analysis of *TAF10KO*^{cEry} and control embryos at E11.5, E12.5 and E13.5. From E12.5 onwards, *TAF10KO*^{cEry} embryos were paler than control or heterozygous littermates and all *TAF10KO*^{cEry} embryos were dead by E13.5. A representative picture of E12.5 embryos is shown in Figure 1A. Gross morphological analysis revealed that the fetal liver size was considerably reduced at E12.5 in *TAF10KO*^{cEry} embryos when compared to control littermates (Figure 1A). The fetal liver size and total blood cell counts were significantly reduced at E13.5 (Figure 1B). Flow cytometry analysis of fetal liver cells showed a decrease in live cells at E13.5 (Figure 1C and Table 1) in concordance with the apoptotic phenotype that has been shown previously in other cell types lacking TAF10(10, 14). This decrease started to be noticeable at E12.5, accompanied by reduced differentiation as measured by flow cytometry (see below). In

summary, erythroid-specific loss of TAF10 dramatically affected the erythroid compartment between E12.5 and E13.5.

We next aimed to determine the differentiation stage at which TAF10 is essential during the fetal erythroid differentiation process**.** Thus, we analyzed E11.5, E12.5 and E13.5 fetal livers by flow cytometry. Despite the fact that gross morphological analysis of E11.5 embryos revealed no significant differences, the distribution of the early erythroid markers KIT and CD71 was altered in *TAF10KO*^{cEry} fetal livers when compared to control littermates, and expression of Ter119 was significantly lower (Table 1). At this stage, the frequency of KIT+ early progenitors was higher and the frequency of KIT+ CD71+ committed erythroid progenitors as well as the more mature CD71+ cell population was reduced (Figure 2A), but not as significant as observed at E12.5 (Table 1). By E12.5, both KIT+ early and KIT+CD71+ committed erythroid progenitors had accumulated at the detriment of the more mature KIT-CD71+ cells. By E13.5 there was an almost complete loss of KIT+CD71+ cells. These data demonstrate a differentiation block of the erythroid progenitors in the fetal liver throughout development. In addition, the KIT MFI (Mean Fluorescence Intensity) was higher in TAF10KO^{cEry} fetal livers by E12.5, although not statistically significant, when compared to control (see Figure 2C and Table 1). The CD71 MFI was reduced at E11.5, and was significantly lower at E12.5 and E13.5 further supporting the notion of a differentiation defect. The maturation delay and the reduction of mature cells (%KIT+, %CD71+, %Ter119+cells, see Table 1) by E12.5 coincided with maximum expression of EpoR-Cre, as measured by the expression of the GFP reporter of EpoR-Cre mice by flow cytometry (Figure 3 and(24)). Furthermore, analysis of CD71 and Ter119 expression, as maturation markers in erythroid cells, revealed a significant reduction in the mature CD71+Ter119+ fraction by E12.5 and a dramatic decrease by E13.5 (Figure 2B). The MFI of Ter119 was reduced throughout development as measured from E11.5 to E13.5 (Figure 2C and Table 1). These data strongly suggest that erythroid development requires TAF10, and that

Molecular and Cellular Biology

erythroid cells are blocked in differentiation upon TAF10 ablation. Since by E13.5 >80% of the fetal liver cells are mature erythroid cells ($CD71^+$ $Ter119^+$) in control embryos (Table 1), it is not surprising that *TAF10KO^{cEry}* embryos do not survive beyond this stage.

RNA-seq analysis of E12.5 *TAF10KO^{cEry}* fetal liver cells shows deregulation of several GATA1 target genes, including *Gata1* itself

To better characterize the TAF10-regulated genes in erythroid cells, we analyzed global gene expression levels by mRNA sequencing on E12.5 fetal liver cells from two homozygous *TAF10KO^{cEry}*, two heterozygous *TAF10KO^{cEry}* and three WT embryos. By using a principal component analysis (PCA) we observed a separate clustering of the homozygous TAF10KO^{cEry} fetal liver samples from the heterozygous TAF10KO^{cEry} and WT controls at E12.5 (Figure 4A). At this stage, when live erythroid cells are still present in the fetal liver, the gene expression analysis revealed around 300 deregulated genes with a minimum of 1.5 fold-change including the *Taf10* gene itself (Figure 4B & 4D). TAF10 levels, as expected, were reduced by more than 50% in *TAF10KO^{cEry}* fetal livers as compared to controls, which corresponds approximately to the percentage of erythroid cells (approx. 50%) beyond the KIT^+ $CD71^+$ committed erythroid progenitor stage in the KO fetal livers (Figure 2A and Table 1), when EpoR expression (and therefore Cre) reaches its maximum levels (as shown above in Figure 3). Gene Ontology analysis (GO) revealed that major 'metabolic pathways', 'cellular response to stress', 'cell type specific apoptosis' and 'cell death' related processes were among the most affected ones (Figure 4C). Expression levels of TFIID subunits were not significantly changed, consistent with previous analysis of TAF10 KO trophoblast and mouse fetal liver(10, 14), suggesting that the composition of the TFIID core complex containing five TAFs (TAF4, 5, 6, 9, 12)(44) does not change upon TAF10 loss. Similarly, the expression of the majority of the SAGA subunits was not affected with the exception of the *Trrap* and *Atxn7l1* genes which showed modest upregulation.

Initially, expression levels of the *Gata1, Gata2, Myb, Klf1,* and *Spi1* genes were measured by qRT-PCR (Figure 4D) and only *Gata1* and *Klf1* displayed a significant change in their expression levels. Subsequently, in the RNA-seq analysis we found that many of the differentially expressed genes are transcription factors and other erythroid-specific genes which have been recently identified as direct GATA1 targets by chromatin immunoprecipitation followed by sequencing (ChIP-seq) using E12.5 mouse fetal liver cells(39). We observed downregulation of genes encoding relevant transcription factors during erythropoiesis, such as *Gata1, Klf1, Nfe2, Zptba7, Bcl2l1* and other metabolically crucial erythroid genes including *Slc4a1, Gypa, Alas2, Car2* (Figure 5A, C). The genes encoding the TFDP22 coregulator, recently reported as an essential factor during terminal erythropoiesis, and the E2F2 transcription factor were significantly downregulated in *TAF10KO*^{cEry} together with two E2F2 target genes (*Dhfr, Ccna2*). We also found a subset of upregulated genes (*Spi1, Gata2, Runx1, Cux1, Car1, Rb1, Cbp/p300, Myc*), most of which are erythroid-related genes (Figure 5B&C). Of note, *Ddit3* and *Trib3,* which were recently reported to be co-induced during erythroid differentiation(45) and the *Ern1* gene, which has been linked to ER stress and induced apoptosis, were found significantly upregulated in the *TAF10* KO embryos (46-48). Important genes for erythroid differentiation, such as *Myb, Tal1, Cdk6* and the recently described *Exosc2* gene, part of the exosome complex (49), did not show any change in their expression levels (q-value >0.4). Most of the aforementioned genes were found to be bound by GATA1 in their regulatory sequences, but in many cases at different developmental stages (39, 45, 49, 50). Specifically for *Myb, Cdk6, and Exosc2 the* GATA1 peaks were found below the cutoff value in E12.5 fetal liver (33).

The expression of globin genes (*Hbb-bh1, Hbb-y, Hbb-bt, Hbb-bs, Hba-a2, Hba-x*) was also lower in the *TAF10KO*^{cEry} fetal liver cells, and especially the *Hbb-y, Hba-a2 and Hba-x* genes were found downregulated with high statistical power (Figure 5A & C). The *Hbb-bh1*

Molecular and Cellular Biology

gene was the only one with no peaks for GATA1 binding at E12.5 among the affected globin genes.

Interestingly, nine genes coding for the DNA-binding proteins (KRAB-ZFPs) shown earlier (51) to be expressed in CD71+TER119+ and/or TER119+ erythroid cells (*Zfp689, Zfp13, Zfp661, Zfp92, Zfp641, Zfp551, Zfp583, Zfp872*) appeared to be unaffected (q-val >0.55) with the exception of *Zfp667* (q-val 0.02) expressed only in TER119+ cells and with no identified GATA1 peak at E12.5, which was downregulated (Figure 5A). In addition, TRIM28 that was recently reported to lead to a block in erythroid differentiation(51) when deleted, did not change in expression, suggesting it is not implicated in the phenotype of the embryos. Most of these genes had no or exceptionally one peak (i.e. *Zfp689*) for GATA1 binding, in their regulatory sequences (±10kb range around the TSS) at E12.5.

These results indicate that TAF10 ablation does not affect global expression as observed before in hepatocyte-specific TAF10 KO cells(10). In addition, about half of the deregulated genes in the *TAF10KO*^{cEry} are potential GATA1 target genes, with many of them known to play a role during erythropoiesis (Figure 5D). In contrast, the expression of genes that have a role in erythroid differentiation, but not bound by GATA1 at this developmental stage (*Zfp689, Zfp13, Myb, Exosc8*), is not affected. Thus, GATA1 targets are among the primary affected genes due to the loss of TAF10 in the *TAF10KO*^{cEry} embryos *in vivo* suggesting an important requirement for a crosstalk or interaction between GATA1 and TAF10.

The composition of TAF10-containing complexes during erythroid differentiation and development

As several recent studies suggested that the composition of general transcription factor or co-activator complexes may change during differentiation and development(22, 23) we sought to analyze the composition of the TAF10-containing TFIID and SAGA complexes

Molecular and Cellular Biology

342 during erythroid differentiation. To this end, protein extracts were prepared from mouse fetal
343 liver cell lines (m^{FLcl}) at the proerythroblast stage (immature), which are able to differentiate
344 to mature erythroblasts in a synchronous manner upon increasing the dose of
345 erythropoietin. TAF10-containing complexes were isolated by anti-TAF10
346 immunoprecipitations (IPs) from immature and differentiating (mature), but still nucleated,
347 erythroid cells and analyzed by mass spectrometry (MS). The relative abundance (emPAI:
348 exponentially modified Protein Abundance Index) of the different subunits in the isolated
349 complexes was first normalized by comparing all abundance values of the subunits of TFIID
350 to TAF1, or all the emPAI values of the subunits of SAGA to TRRAP, the largest subunits in
351 each respective complex (Figure 6A&B). Interestingly, when comparing the composition of
352 the TFIID complexes between proerythroblasts and mature erythroblasts, we noticed that
353 TAF4b completely disappeared from the TFIID complex of mature cells and also less TAF4
354 was associated with the TFIID complex in mature erythroid cells. In agreement with the
355 observation that TAF12 is the histone-fold partner of TAF4(52, 53), the TAF4/TAF4b
356 decrease in mature erythroid TFIID was accompanied by a reduction of TAF12. Note that
357 TAF9, but not TAF9b, was also significantly decreased in TAF10 IPs from mature erythroid
358 cell extracts. When comparing the composition of the SAGA complex at these two
359 differentiation stages after normalizing to TRRAP, we observed a slight reduction in several
360 subunits of SAGA complex in TAF10 IPs from mature erythroid cell extracts (Figure 6B).
361 Out of the two homologous SAGA HATs (GCN5 and PCAF), which have been reported to
362 be mutually exclusive in the corresponding SAGA complexes(54), the abundance of GCN5
363 is reduced about three times, while that of PCAF is not reduced, in mature erythroid cells.
364 Moreover, in the deubiquitination (DUB) module of SAGA, ATXN7 and its orthologue,
365 ATXN7L2 appear to be replaced by the orthologous ATXN7L1 at the mature stage. We note
366 that the absence of TAF13 from TFIID and of ENY2 from SAGA in these analyses may be
367 due to the very small size of these proteins.

Next, we sought to analyze the composition of human TAF10-containing complexes during development (i.e. fetal and adult stages). Although the erythroid cells are at the proerythroblast stage at both developmentally different niches (fetal and adult) and should not be directly compared to the mouse situation we have made similar observations regarding the slight stoichiometric changes that occur in the corresponding subunits of TFIID and SAGA. Our analyses demonstrate that all of the subunits of the two complexes are present in the corresponding complexes at these two developmental stages (Figure 6C, D).

Importantly, these results together indicate that the canonical composition of the two analyzed TAF10-containing complexes, TFIID and SAGA, does not change dramatically during mouse and human erythroid differentiation and development. However, especially in mouse complexes we observed often significant stoichiometric changes of those specific subunits that have orthologues (TAF4/TAF4b, GCN5/PCAF, TAF9/TAF9b, ATXN7/ATXN7L1/ATXNL2), which may slightly affect the function of these transcription complexes.

TAF10 and GATA1 interact in mouse and human fetal liver cells

As our TAF10 KO experiments suggested a functional crosstalk between GATA1 and TAF10, we next analyzed whether specific erythroid transcription factors, such as GATA1, would co-immunoprecipitate with TAF10 from human fetal liver or human peripheral blood erythroid progenitor cultures and mouse fetal liver cell lines. Importantly, GATA1, was identified by mass spectrometry (MS) as a TAF10 interactor in h^{FL} cells together with other previously reported activators and cofactors (Table 2), such as LDB1 and TAL1, which are components of the so-called pentameric complex(55).

We also analyzed GATA1 interactors by carrying out an anti-GATA1 IP and subsequent MS in MEL cells. Endogenous TAF10 together with other TAFs and SAGA

Molecular and Cellular Biology

394 subunits were identified in the MS analysis (Table 3). Consequently, immunoprecipitation of
395 GATA1 from MEL cells revealed that endogenous TAF10 co-immunoprecipitated with
396 GATA1 and the Friend of GATA1 cofactor (FOG1) (Figure 7A). In addition, the interaction of
397 endogenous TAF10 and GATA1 was verified by immunoprecipitating TAF10 from nuclear
398 extracts prepared from MEL cells and the co-immunoprecipitated GATA1 was analyzed by
399 Western blotting (Figure 7B). Similar IP assays were performed with E12.5 fetal liver cells or
400 MEL cells that express biotinylated GATA1 (bio-GATA1) (Figure 8). These experiments
401 further confirmed the interaction between TAF10-containing complexes and GATA1
402 previously identified in the MS data.

403 We further investigated whether the GATA1-TAF10 interaction is direct by *in vitro*
404 protein-protein interaction experiments. First either TAF10 alone, or TAF8-TAF10
405 heterodimer was immune-purified using an anti-TAF10 antibody from SFf9 extracts, in
406 which the corresponding proteins were overexpressed using the baculovirus system. Next,
407 purified GST-GATA1 protein or GST alone, were added to TAF10 or TAF10-TAF8 bound
408 beads (Figure 7C). After several washing steps with high salt buffer, bead-bound proteins
409 were denatured, resolved on a SDS-PAGE and analyzed by Western blotting using an anti-
410 GST antibody. Our *in vitro* results indicate that GST-GATA1 bound to both TAF10- and
411 TAF8-TAF10-containing beads, but GST alone did not (Figure 7D). These results together
412 support a direct interaction between the key erythroid transcription factor GATA1 and
413 TAF10.

414

415 **TAF10 is bound to GATA1 sites in the GATA1 locus**

416 In light of the downregulation of GATA1 mRNA levels we wished to determine whether
417 TAF10 would be present, and thus possibly regulate the expression of the *GATA1* gene.
418 We looked specifically at a palindromic GATA1 binding site known to be required for normal
419 *GATA1* transcription, and an additional GATA1 binding site, which locates next to a TATA-

box at -3kb relative to the TSS(56, 57) (Figure 9A). We performed chromatin immunoprecipitation (ChIP) of TAF10 and GATA1 in h^{FL} and h^{BL} cells representing two distinct developmental stages as defined by the expression of fetal and adult hemoglobin respectively. We found that TAF10 and GATA1 binding was more enriched at both GATA1 binding sites at the human GATA1 locus in the fetal liver as compared to adult blood proerythroblasts (Figure 9B-C). However, while GATA1 binds at both binding sites examined at both developmental stages, TAF10 is clearly not bound at the palindromic GATA1 binding site in the adult stage and is significantly less enriched than GATA1 at the -3kb binding site in the fetal stage (Figure 9C).

GATA1 auto-regulates its expression by binding to its own promoter and enhancers. We detected TAF10-GATA1 protein-protein interaction mainly in h^{FL} extracts and ChIP results show the selective binding of TAF10 at the palindromic GATA1 site during fetal stages. Collectively, these results support the notion that TAF10 has a role in the developmental regulation of GATA1 expression.

Molecular and Cellular Biology

Molecular and Cellular Biology

Discussion

Erythropoiesis is a process that is controlled tightly by the regulated expression of erythroid specific transcription factors and their interactions with cofactors. General transcription factors also have an active role, i.e. they have a more cell-type specific function than it was originally thought(58). Whether the TAF10 component of TFIID and SAGA exerts such a role by shaping the interactions with activators in erythroid differentiation and development was the topic of this study.

TAF10 was specifically ablated to disrupt the canonical TFIID and SAGA complexes in erythroid cells from early stages of mouse development (E8.0) in mice by crossing *Taf10Lox* with *EpoR-Cre* mice. This resulted in a block in erythropoiesis leading to embryonic death at around E13.5. A progressive delay in the differentiation kinetics through development, which starts already at E11.5, is more pronounced at E12.5 with an accumulation of CD71+TER119+ cells at the expense of mature TER119+ erythroid cells. Expression analysis by mRNA sequencing (RNA-seq) showed that erythroid transcription factor genes *Gata1, Klf1, Nfe2, Zbtb7a, and Bcl2l1* were within the downregulated group of genes and expression of *Gata2, Spi1* and *Myb* did not change significantly, while *Runx1, Rb1, Myc* and *Cited2* were upregulated following the opposite gene expression pattern that characterizes the transition from the CD71+TER119+ towards to TER119+ mature erythroid cells (59). Along the same lines, the TFDP2 coregulator and E2F2 transcription factor were also found downregulated, and both genes were under the regulatory control of GATA1. Normally, these two genes are induced upon erythroid differentiation and it was shown that downregulation of *Tfdp2* prevents proper erythroid differentiation(45). Although *E2f2* was downregulated significantly, the expression level of several E2F2 target genes was reduced (*Dhfr, Ccna2*) in contrast to what we would expect due to the repressive role(45) of E2F2 on its targets. Globin genes were also downregulated as a clear sign of the differentiation block during erythroid differentiation. In contrast, genes with less well defined roles in the

erythroid lineage and no peaks for GATA1 in the ChIP-seq analysis at this developmental stage (E12.5), i.e. KRAB-ZFP proteins, did not change their expression levels in the *TAF10KO*^{cEry} fetal livers.

Nine of these KRAB-ZFP proteins are known to be expressed in erythroid progenitors and two of them are required for proper erythroid differentiation (ZFP689, ZFP13)(51). These results suggest that TAF10 is required to stabilize the PIC complex at certain loci, explaining why the transcription of specific genes is affected upon TAF10 loss. This finding is in agreement with the notion that TAF10 loss does not affect general transcription as would be expected for a cornerstone TAF of the TFIID(44). The specific gene expression changes in TAF10 KO erythroid cells could explain their progressive block in differentiation and subsequent apoptosis as identified by our GO analysis in *TAF10KO*^{cEry} fetal livers, since GATA1 KO erythroid cells are unable to differentiate but arrest their cell cycle and undergo apoptosis. In addition, three genes linked to ER stress and induced apoptosis (*Ddit3*(46), *Trib3*(60) and *Ern1*(48)), were found significantly upregulated in the *TAF10KO*^{cEry} fetal livers supporting the notion that apoptosis is among the main causes of the observed phenotype around E13.5.

In parallel with the *in vivo* studies we defined the composition of TAF10-containing complexes in erythroid cells in order to investigate the dynamic changes that appear to be crucial in other cell types during differentiation(61). We performed MS analysis of human and mouse cultured erythroid progenitor cells at an immature stage (proerythroblasts) and upon differentiation, which revealed that most of the TFIID and SAGA subunits are present in these complexes at all stages analyzed. Our data exclude a total rearrangement of the TFIID or SAGA complexes in this differentiation system, in contrast to what was reported for TFIID during liver hepatocyte(62) or myoblast differentiation(63). Nevertheless, we observe a dynamic reorganization of some TFIID and SAGA subunits, ususally affecting those that have paralogues, such as TAF4/TAF4b, TAF9/TAF9b, GCN5/PCAF and

Molecular and Cellular Biology

ATXN7/ATXN7L1/ATXN7L2, during differentiation without affecting the core structure of these complexes. Interestingly, Pijnappel et al(64) demonstrated that overexpression of TAF4 with the pluripotency factors, and presumably the incorporation of TAF4 into pre-existing TFIID complexes lacking TAF4, can efficiently reprogram differentiated cells into induced pluripotent stem cells (iPSCs). The TFIID complexes we purified from mouse erythroid differentiated cells also have reduced TAF4/TAF12 heterodimers and no TAF4b, whereas TAF4/TAF12 heterodimers and TAF4b are present at the immature stages of differentiation. This is in excellent agreement with the results of the Timmers lab (51) and suggests that the low level of TAF4 and/or the lack of TAF4b is associated with a differentiated state, whereas immature stage cells contain a TFIID complex with stoichiometric amounts of TAF4/4b.

The idea that developmental gene regulation is dependent on protein interactions between TFIID, activators and co-activators is also supported by our results. GATA1 and its well known partners LDB1 and TAL1 were found to interact with endogenous TAF10 in the fetal liver cells of both mouse and human origin. This interaction was verified by reciprocal immunoprecipitation (anti-TAF10 and anti-GATA1 IPs) in MEL cells and *in vitro* by using purified GST-GATA1 and TAF10 (or TAF10-TAF8 heterodimers) proteins. Similar interactions have been reported for activation of the β-globin gene between KLF1 and TAF9(65) as well as GATA1 and MED1(66, 67) in erythroid cells.

Other TFs and co-factors, including subunits of the CCR4-NOT complex (e.g CNOT3), CBX3 (a paralogue of CBX1 (HP1β)) and TRIM28 were also found to interact with TAF10 by mass spectrometry (Table 2). CNOT3, CBX1 and TRIM28 were previously reported to form a unique module involved in developmental processes(68) and TRIM28 in particular has an important role in erythropoiesis(51, 69). Of note, CNOT3 and TRIM28 do not physically interact, while there have been reports of interactions of TFIID with the CCR4-NOT complex(70). Therefore, TFIID might be acting as a scaffold protein for the assembly

Molecular and Cellular Biology

513 of the module and might play an important role in the developmental processes controlled
514 by this module, including erythropoiesis, which would be interesting to investigate in the
515 future.

516 Interestingly, our observation that TAF10 binding was enriched at the promoter of the
517 human GATA1 locus in fetal erythroid cells as compared to adult erythroid cells together
518 with the observed TAF10-GATA1 interaction suggest that there is a role for TAF10 in the
519 regulation of GATA1 transcription, which contributes to the phenotype observed in
520 *TAF10KO*^{cEry} embryos. This interaction preferentially occurs during the fetal stages of
521 erythropoiesis, indicating that it is a developmental-specific event exerting its effect mainly,
522 but not exclusively, on GATA1 target genes as observed in the RNA-seq data of the mouse
523 fetal livers. We know that GATA1 expression levels do not change in human fetal and adult
524 erythroblasts(71). However, there are dynamic changes in the occupancy of transcription
525 factors and consequently in protein-protein interactions, involving master regulators, that
526 could potentially activate, repress(72) or stabilize gene expression levels. When such
527 interactions are disturbed, transcriptional deregulation is not global but it depends on the
528 transcriptional state of the gene at that developmental stage(10).

529 We propose that TFIID and SAGA contribute to this dynamic landscape of
530 developmental-specific protein interactions through TAF10 interaction with GATA1 thus
531 contributing to development and differentiation of the erythroid lineage.

532

Acknowledgements

The authors would like to thank the PRIDE team for their help and support regarding the deposition of the mass spectrometry data. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository [1] with the dataset identifier PXD000729 and DOI 10.6019/PXD000729. The GST and GST-GATA1 plasmids were a generous gift from K. Freson at KU Leuven (Molecular and Vascular Biology). We are also grateful to the members of T. Economou lab (Molecular Bacteriology) at KU Leuven and especially to N. Famelis and K. Tsolis for helping with the purification of the GST fusion proteins and Petros Kolovos from Erasmus MC for technical support.

This work has been supported partially by an EMBO short-term fellowship ASTF 15-2010 (P.P.), the Netherlands Organization for Scientific Research (NWO-VENI 863.09.012, L.G.), the Netherlands Genomics Initiative (NGI Zenith 93511036), the Netherlands Proteomics Center (NPC), the Netherlands Initiative of Regenerative Medicine (NIRM) and the EU Integrated project EuTRACC to F.G. and L.T.; the Landsteiner Foundation for Blood Transfusion Research (LSBR; 1040) to FG and SP; ZonMw (TOP 40-00812-98-12128 and DN 82-301), EU FP7 Specific Cooperation Research Project “THALAMOSS” (306201) to S.P. and ERC Advanced (Birtoaction, N° 340551) grant to L.T.

Contribution**:** P.P, F.G., and L.T. designed the study; P.P., L.G., J.D., D.N.P., E.K., R.vdL., F.P., E.S., performed the experiments; P.P., L.G., J.D., H.J.G.vdW., D.H.W.D., P.V., J.S., S.P., F.G., L.T. analyzed the data; P.P., L.G., S.P., F.G., L.T. wrote the paper.

References

1. **Muller F, Tora L.** 2013. Chromatin and DNA sequences in defining promoters for transcription initiation. Biochimica et biophysica acta doi:10.1016/j.bbagrm.2013.11.003.
2. **Gershenzon NI, Ioshikhes IP.** 2005. Synergy of human Pol II core promoter elements revealed by statistical sequence analysis. Bioinformatics **21:**1295-1300.
3. **Zhang MQ.** 1998. Identification of human gene core promoters in silico. Genome research **8:**319-326.
4. **Gangloff YG, Romier C, Thuault S, Werten S, Davidson I.** 2001. The histone fold is a key structural motif of transcription factor TFIID. Trends in biochemical sciences **26:**250-257.
5. **Kaufmann J, Smale ST.** 1994. Direct recognition of initiator elements by a component of the transcription factor IID complex. Genes Dev **8:**821-829.
6. **Chalkley GE, Verrijzer CP.** 1999. DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. EMBO J **18:**4835-4845.
7. **Juven-Gershon T, Kadonaga JT.** 2010. Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev Biol **339:**225-229.
8. **Vassallo MF, Tanese N.** 2002. Isoform-specific interaction of HP1 with human TAFII130. Proc Natl Acad Sci U S A **99:**5919-5924.
9. **Liu WL, Coleman RA, Ma E, Grob P, Yang JL, Zhang Y, Dailey G, Nogales E, Tjian R.** 2009. Structures of three distinct activator-TFIID complexes. Genes Dev **23:**1510-1521.
10. **Tatarakis A, Margaritis T, Martinez-Jimenez CP, Kouskouti A, Mohan WS, 2nd, Haroniti A, Kafetzopoulos D, Tora L, Talianidis I.** 2008. Dominant and redundant functions of TFIID involved in the regulation of hepatic genes. Mol Cell **31:**531-543.
11. **Frontini M, Soutoglou E, Argentini M, Bole-Feysot C, Jost B, Scheer E, Tora L.** 2005. TAF9b (formerly TAF9L) is a bona fide TAF that has unique and overlapping roles with TAF9. Mol Cell Biol **25:**4638-4649.
12. **Spedale G, Timmers HT, Pijnappel WW.** 2012. ATAC-king the complexity of SAGA during evolution. Genes & development **26:**527-541.
13. **Timmers HT, Tora L.** 2005. SAGA unveiled. Trends Biochem Sci **30:**7-10.
14. **Mohan WS, Jr., Scheer E, Wendling O, Metzger D, Tora L.** 2003. TAF10 (TAF(II)30) is necessary for TFIID stability and early embryogenesis in mice. Mol Cell Biol **23:**4307-4318.
15. **Indra AK, Mohan WS, 2nd, Frontini M, Scheer E, Messaddeq N, Metzger D, Tora L.** 2005. TAF10 is required for the establishment of skin barrier function in foetal, but not in adult mouse epidermis. Dev Biol **285:**28-37.
16. **Wong PM, Chung SW, Eaves CJ, Chui DH.** 1985. Ontogeny of the mouse hemopoietic system. Prog Clin Biol Res **193:**17-28.
17. **Sankaran VG, Xu J, Orkin SH.** 2010. Advances in the understanding of haemoglobin switching. Br J Haematol **149:**181-194.
18. **Ferreira R, Ohneda K, Yamamoto M, Philipsen S.** 2005. GATA1 function, a paradigm for transcription factors in hematopoiesis. Mol Cell Biol **25:**1215-1227.
19. **Gutierrez L, Nikolic T, van Dijk TB, Hammad H, Vos N, Willart M, Grosveld F, Philipsen S, Lambrecht BN.** 2007. Gata1 regulates dendritic-cell development and survival. Blood **110:**1933-1941.
20. **Fujiwara Y, Browne CP, Cunniff K, Goff SC, Orkin SH.** 1996. Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. Proc Natl Acad Sci U S A **93:**12355-12358.
21. **Gutierrez L, Tsukamoto S, Suzuki M, Yamamoto-Mukai H, Yamamoto M, Philipsen S, Ohneda K.** 2008. Ablation of Gata1 in adult mice results in aplastic crisis, revealing its essential role in steady-state and stress erythropoiesis. Blood **111:**4375-4385.
22. **Muller F, Zaucker A, Tora L.** 2010. Developmental regulation of transcription initiation: more than just changing the actors. Curr Opin Genet Dev **20:**533-540.

23. **Goodrich JA, Tjian R.** 2010. Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. Nat Rev Genet **11:**549-558.
24. **Heinrich AC, Pelanda R, Klingmuller U.** 2004. A mouse model for visualization and conditional mutations in the erythroid lineage. Blood **104:**659-666.
25. **Tsai SF, Strauss E, Orkin SH.** 1991. Functional analysis and in vivo footprinting implicate the erythroid transcription factor GATA-1 as a positive regulator of its own promoter. Genes Dev **5:**919-931.
26. **Demeny MA, Soutoglou E, Nagy Z, Scheer E, Janoshazi A, Richardot M, Argentini M, Kessler P, Tora L.** 2007. Identification of a small TAF complex and its role in the assembly of TAF-containing complexes. PLoS One **2:**e316.
27. **Jacq X, Brou C, Lutz Y, Davidson I, Chambon P, Tora L.** 1994. Human TAFII30 is present in a distinct TFIID complex and is required for transcriptional activation by the estrogen receptor. Cell **79:**107-117.
28. **Wieczorek E, Brand M, Jacq X, Tora L.** 1998. Function of TAF(II)-containing complex without TBP in transcription by RNA polymerase II. Nature **393:**187-191.
29. **Soutoglou E, Demeny MA, Scheer E, Fienga G, Sassone-Corsi P, Tora L.** 2005. The nuclear import of TAF10 is regulated by one of its three histone fold domain-containing interaction partners. Mol Cell Biol **25:**4092-4104.
30. **van den Berg DL, Snoek T, Mullin NP, Yates A, Bezstarosti K, Demmers J, Chambers I, Poot RA.** 2010. An Oct4-centered protein interaction network in embryonic stem cells. Cell stem cell **6:**369-381.
31. **Leberbauer C, Boulme F, Unfried G, Huber J, Beug H, Mullner EW.** 2005. Different steroids co-regulate long-term expansion versus terminal differentiation in primary human erythroid progenitors. Blood **105:**85-94.
32. **Follows GA, Tagoh H, Lefevre P, Hodge D, Morgan GJ, Bonifer C.** 2003. Epigenetic consequences of AML1-ETO action at the human c-FMS locus. EMBO J **22:**2798-2809.
33. **Schmittgen TD, Livak KJ.** 2008. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc **3:**1101-1108.
34. **Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S, Gil L, Giron CG, Gordon L, Hourlier T, Hunt S, Johnson N, Juettemann T, Kahari AK, Keenan S, Kulesha E, Martin FJ, Maurel T, McLaren WM, Murphy DN, Nag R, Overduin B, Pignatelli M, Pritchard B, Pritchard E, Riat HS, Ruffier M, Sheppard D, Taylor K, Thormann A, Trevanion SJ, Vullo A, Wilder SP, Wilson M, Zadissa A, Aken BL, Birney E, Cunningham F, Harrow J, Herrero J, Hubbard TJ, Kinsella R, Muffato M, Parker A, Spudich G, Yates A, et al.** 2014. Ensembl 2014. Nucleic acids research **42:**D749-755.
35. **Anders S, Pyl PT, Huber W.** 2014. HTSeq - A Python framework to work with high-throughput sequencing data. Bioinformatics doi:10.1093/bioinformatics/btu638.
36. **Anders S, Huber W.** 2010. Differential expression analysis for sequence count data. Genome Biol **11:**R106.
37. **Hochberg Y, Benjamini Y.** 1990. More powerful procedures for multiple significance testing. Statistics in medicine **9:**811-818.
38. **Falcon S, Gentleman R.** 2007. Using GOstats to test gene lists for GO term association. Bioinformatics **23:**257-258.
39. **Papadopoulos GL, Karkoulia E, Tsamardinos I, Porcher C, Ragoussis J, Bungert J, Strouboulis J.** 2013. GATA-1 genome-wide occupancy associates with distinct epigenetic profiles in mouse fetal liver erythropoiesis. Nucleic acids research **41:**4938-4948.
40. **von Lindern M, Deiner EM, Dolznig H, Parren-Van Amelsvoort M, Hayman MJ, Mullner EW, Beug H.** 2001. Leukemic transformation of normal murine erythroid progenitors: v- and c-ErbB act through signaling pathways activated by the EpoR and c-Kit in stress erythropoiesis. Oncogene **20:**3651-3664.

41. **Lee R, Kertesz N, Joseph SB, Jegalian A, Wu H.** 2001. Erythropoietin (Epo) and EpoR expression and 2 waves of erythropoiesis. Blood **98:**1408-1415.
42. **Isern J, Fraser ST, He Z, Baron MH.** 2010. Developmental niches for embryonic erythroid cells. Blood Cells Mol Dis **44:**207-208.
43. **Palis J.** 2014. Primitive and definitive erythropoiesis in mammals. Front Physiol **5:**3.
44. **Bieniossek C, Papai G, Schaffitzel C, Garzoni F, Chaillet M, Scheer E, Papadopoulos P, Tora L, Schultz P, Berger I.** 2013. The architecture of human general transcription factor TFIID core complex. Nature **493:**699-702.
45. **Chen C, Lodish HF.** 2014. Global analysis of induced transcription factors and cofactors identifies Tfdp2 as an essential coregulator during terminal erythropoiesis. Exp Hematol **42:**464-476 e465.
46. **Lu M, Lawrence DA, Marsters S, Acosta-Alvear D, Kimmig P, Mendez AS, Paton AW, Paton JC, Walter P, Ashkenazi A.** 2014. Cell death. Opposing unfolded-protein-response signals converge on death receptor 5 to control apoptosis. Science **345:**98-101.
47. **Harding HP, Zhang Y, Ron D.** 1999. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. Nature **397:**271-274.
48. **Brewer JW, Hendershot LM, Sherr CJ, Diehl JA.** 1999. Mammalian unfolded protein response inhibits cyclin D1 translation and cell-cycle progression. Proc Natl Acad Sci U S A **96:**8505-8510.
49. **McIver SC, Kang YA, DeVilbiss AW, O'Driscoll CA, Ouellette JN, Pope NJ, Camprecios G, Chang CJ, Yang D, Bouhassira EE, Ghaffari S, Bresnick EH.** 2014. The exosome complex establishes a barricade to erythroid maturation. Blood **124:**2285-2297.
50. **Yu M, Riva L, Xie H, Schindler Y, Moran TB, Cheng Y, Yu D, Hardison R, Weiss MJ, Orkin SH, Bernstein BE, Fraenkel E, Cantor AB.** 2009. Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. Mol Cell **36:**682-695.
51. **Barde I, Rauwel B, Marin-Florez RM, Corsinotti A, Laurenti E, Verp S, Offner S, Marquis J, Kapopoulou A, Vanicek J, Trono D.** 2013. A KRAB/KAP1-miRNA cascade regulates erythropoiesis through stage-specific control of mitophagy. Science **340:**350-353.
52. **Gangloff YG, Werten S, Romier C, Carre L, Poch O, Moras D, Davidson I.** 2000. The human TFIID components TAF(II)135 and TAF(II)20 and the yeast SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like pairs. Molecular and cellular biology **20:**340-351.
53. **Leurent C, Sanders S, Ruhlmann C, Mallouh V, Weil PA, Kirschner DB, Tora L, Schultz P.** 2002. Mapping histone fold TAFs within yeast TFIID. EMBO J **21:**3424-3433.
54. **Nagy Z, Riss A, Fujiyama S, Krebs A, Orpinell M, Jansen P, Cohen A, Stunnenberg HG, Kato S, Tora L.** 2010. The metazoan ATAC and SAGA coactivator HAT complexes regulate different sets of inducible target genes. Cellular and molecular life sciences : CMLS **67:**611-628.
55. **Soler E, Andrieu-Soler C, de Boer E, Bryne JC, Thongjuea S, Stadhouders R, Palstra RJ, Stevens M, Kockx C, van Ijcken W, Hou J, Steinhoff C, Rijkers E, Lenhard B, Grosveld F.** 2010. The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. Genes Dev **24:**277-289.
56. **Nicolis S, Bertini C, Ronchi A, Crotta S, Lanfranco L, Moroni E, Giglioni B, Ottolenghi S.** 1991. An erythroid specific enhancer upstream to the gene encoding the cell-type specific transcription factor GATA-1. Nucleic Acids Res **19:**5285-5291.
57. **Moriguchi T, Yu L, Takai J, Hayashi M, Satoh H, Suzuki M, Ohneda K, Yamamoto M.** 2015. The Human GATA1 Gene Retains a 5' Insulator that Maintains Chromosomal Architecture and GATA1 Expression Levels in Splenic Erythroblasts. Mol Cell Biol doi:10.1128/MCB.00011-15.
58. **Naar AM, Lemon BD, Tjian R.** 2001. Transcriptional coactivator complexes. Annual review of biochemistry **70:**475-501.

59. **Wong P, Hattangadi SM, Cheng AW, Frampton GM, Young RA, Lodish HF.** 2011. Gene induction and repression during terminal erythropoiesis are mediated by distinct epigenetic changes. Blood **118:**e128-138.
60. **Ohoka N, Yoshii S, Hattori T, Onozaki K, Hayashi H.** 2005. TRB3, a novel ER stress-inducible gene, is induced via ATF4-CHOP pathway and is involved in cell death. EMBO J **24:**1243-1255.
61. **Liu Z, Scannell DR, Eisen MB, Tjian R.** 2011. Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. Cell **146:**720-731.
62. **D'Alessio JA, Ng R, Willenbring H, Tjian R.** 2011. Core promoter recognition complex changes accompany liver development. Proc Natl Acad Sci U S A **108:**3906-3911.
63. **Deato MD, Tjian R.** 2007. Switching of the core transcription machinery during myogenesis. Genes & development **21:**2137-2149.
64. **Pijnappel WW, Esch D, Baltissen MP, Wu G, Mischerikow N, Bergsma AJ, van der Wal E, Han DW, Bruch H, Moritz S, Lijnzaad P, Altelaar AF, Sameith K, Zaehres H, Heck AJ, Holstege FC, Scholer HR, Timmers HT.** 2013. A central role for TFIID in the pluripotent transcription circuitry. Nature **495:**516-519.
65. **Sengupta T, Cohet N, Morle F, Bieker JJ.** 2009. Distinct modes of gene regulation by a cell-specific transcriptional activator. Proc Natl Acad Sci U S A **106:**4213-4218.
66. **Stumpf M, Waskow C, Krotschel M, van Essen D, Rodriguez P, Zhang X, Guyot B, Roeder RG, Borggrefe T.** 2006. The mediator complex functions as a coactivator for GATA-1 in erythropoiesis via subunit Med1/TRAP220. Proc Natl Acad Sci U S A **103:**18504-18509.
67. **Stumpf M, Yue X, Schmitz S, Luche H, Reddy JK, Borggrefe T.** 2010. Specific erythroid-lineage defect in mice conditionally deficient for Mediator subunit Med1. Proc Natl Acad Sci U S A **107:**21541-21546.
68. **Hu G, Kim J, Xu Q, Leng Y, Orkin SH, Elledge SJ.** 2009. A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes Dev **23:**837-848.
69. **Hosoya T, Clifford M, Losson R, Tanabe O, Engel JD.** 2013. TRIM28 is essential for erythroblast differentiation in the mouse. Blood **122:**3798-3807.
70. **Collart MA, Timmers HT.** 2004. The eukaryotic Ccr4-not complex: a regulatory platform integrating mRNA metabolism with cellular signaling pathways? Prog Nucleic Acid Res Mol Biol **77:**289-322.
71. **Xu J, Shao Z, Glass K, Bauer DE, Pinello L, Van Handel B, Hou S, Stamatoyannopoulos JA, Mikkola HK, Yuan GC, Orkin SH.** 2012. Combinatorial assembly of developmental stage-specific enhancers controls gene expression programs during human erythropoiesis. Developmental cell **23:**796-811.
72. **Pimkin M, Kossenkov AV, Mishra T, Morrissey CS, Wu W, Keller CA, Blobel GA, Lee D, Beer MA, Hardison RC, Weiss MJ.** 2014. Divergent functions of hematopoietic transcription factors in lineage priming and differentiation during erythro-megakaryopoiesis. Genome research doi:10.1101/gr.164178.113.

751 **Figure legends**

752

753 **Figure 1. Developmental analysis of *TAF10KO^{cEry}***

754 **A.** Representative pictures of control and *TAF10KO^{cEry}* embryos at E12.5. *TAF10KO^{cEry}*
755 embryos are pale and the fetal liver is much smaller compared to those of the control
756 embryos.

757 **B.** Cell count numbers at E13.5 from fetal liver and embryonic blood of control and
758 *TAF10KO^{cEry}* embryos.

759 **C.** Flow cytometry analysis of fetal liver single cell suspensions of control and *TAF10KO^{cEry}*
760 embryos at different developmental stages (E11.5, E12.5 and E13.5). Numbers indicate the
761 percentages of live cells (Hoechst negative).

762

763 **Figure 2. Flow cytometry of *TAF10KO^{cEry}* fetal liver cells during gestation**

764 Representative flow cytometry analysis of fetal liver single cell suspension of control and
765 *TAF10KO^{cEry}* embryos at E11.5, E12.5 and E13.5 following the differentiation of erythroid
766 cells.

767 **A.** Staining with KIT and CD71 markers (numbers indicate percentages of gated live cells).

768 **B.** Staining with CD71 and Ter119 markers (numbers indicate percentages of gated live
769 cells).

770 **C.** MFI (Mean Fluorescence Intensity) graphs of KIT, CD71 and Ter119 populations as
771 expressed during development of erythroid cells of control and *TAF10KO^{cEry}* fetal liver cells.

772

773

774 **Figure 3. Analysis of *EpoR-Cre* transgenic expression**

775 Flow cytometry analysis of fetal livers cells following the expression of the GFP protein
776 expressed in frame with Cre recombinase in *EpoR-Cre* mice. Expression of GFP confined

Molecular and Cellular Biology

777 to the erythroid compartment was detected initially in KIT+/CD71+ cells at E11.5 (green
778 population for controls and red population for *TAF10KO*^{cEry} erythroid cells). Erythroid cells
779 are affected in *TAF10KO*^{cEry} embryos and that is why the percentage (numbers shown in
780 the graphs) of CD71+/GFP+ cells is lower as compared to control cells at E12.5. By E13.5
781 most of the cells are dead.
782
783 **Figure 4. Gene Expression analysis of E12.5 *TAF10KO*^{cEry} fetal livers**
784 **A.** Principal Component analysis (PCA) plot of the first two components of the RNA-seq
785 fetal liver samples (*TAF10KO*^{cEry}, Heterozygous *TAF10KO*, WT). The variance explained by
786 each component is depicted between parenthesis on the axes.
787 **B.** MA-plot of the mean normalized gene count versus the $\log_2$ fold changes of
788 *TAF10KO*^{cEry} over WT and Heterozygous *TAF10KO*. Genes are plotted as closed black
789 circles. Genes with an adjusted *p*-value (FDR) < 0.01 are colored red. Genes that fall out of
790 the window boundaries of -2 or 2 $\log_2$ fold change are plotted as open triangles.
791 **C.** Gene ontology (GO) analysis of the upregulated genes in the *TAF10KO*^{cEry} fetal livers.
792 Metabolic pathways, apoptotic and cell death related processes and proliferation are among
793 the most affected.
794 **D.** qRT-PCR on total mRNA of fetal liver cells at E12.5. Expression levels of transcription
795 factors and globin genes are depicted. Bars represent standard error of the mean (SEM).
796
797 **Figure 5. Gene Expression analysis (Transcription factors, erythroid-related and**
798 **globin genes)**
799 **A.** RNA-seq gene expression levels (FPKM: **F**ragments **P**er **K**ilobase of exon per **M**illion
800 fragments mapped) of upregulated erythroid-related genes with a q-value<0.05 and GATA1
801 binding peaks found within ±10Kb of their TSS ,
802 **B.** Downregulated genes with a q-value<0.05 with GATA1 binding peaks found,

C. Deregulated genes with a q-value>0.05 with GATA1 binding peaks found, and

D. Venn diagram of deregulated genes (upreg: upregulated in *TAF10KO*^{cEry} fetal liver cells-green circle, downreg: downregulated genes in *TAF10KO*^{cEry} fetal liver cells-purple circle) as identified by RNA-seq analysis (q-value <0.05) and GATA1 target genes described in (39)(red circle). Half of the deregulated genes have at least one GATA1 binding peak at ±10Kb of their TSS.

Figure 6. Mass spectrometry of TFIID and SAGA complexes in mouse and human erythroid cells

A and B. TAF10-containing complexes, TFIID (in A) and SAGA (in B), were isolated by a TAF10 immunoprecipitation from protein extracts prepared from immature and differentiating (mature) mouse erythroid cells and analyzed by mass spectrometry. The relative abundance (emPAI: exponentially modified Protein Abundance Index) values of the different subunits in the isolated complexes were first normalized by comparing the abundance values of all the subunits of TFIID to TAF1, or all the subunits of SAGA to TRRAP, the largest subunits in each complex.

C and D. Similarly, TAF10-containing complexes, TFIID (in C) and SAGA (in D), were isolated by a TAF10 immunoprecipitation from protein extracts prepared from human fetal liver (hFL) and human peripheral blood erythroid progenitors (hBL) and analyzed by mass spectrometry. The emPAI values presented were normalized as described for the mouse samples.

Figure 7. Immunoprecipitation of TAF10 and GATA1 in MEL cells

A. GATA1 immunoprecipitation in MEL cells. Anti-GATA1 antibody (N6) was used for Western blot analysis. TAF10 and FOG-1 are co-immunoprecipitated.

B. TAF10 immunoprecipitation in MEL nuclear extracts using the 6TA 2B11 antibody clone. IgG antibody was used as a control. GATA1 is detected in the IP fraction, confirming the MS results. Sup: supernatant.

C. TAF10 alone, or TAF8-TAF10 heterodimer was immune-purified using an anti-TAF10 antibody from SFf9 extracts and tested by Western blot using the indicated antibodies. GST-GATA1 protein, or GST, were also purified and tested by coomassie blue (CBB) staining.

D. The purified proteins were combined as indicated, incubated and after several washes the beads-bound proteins were denatured, resolved on a SDS-PAGE and analyzed by Western blot using an anti- GST antibody. Antibody heavy (AbHc) and light (AbLc) chains are indicated.

Figure 8. **Immunoprecipitations of TAF10 and GATA1 in MEL cells and mouse fetal liver E12.5**

A. MEL cells and mouse fetal liver nuclear cell extracts were used to immunoprecipitate GATA1 endogenous protein or the bio-GATA1 over-expressed (O/E) in MEL cells. GATA1 (N6) IP or streptavidin pulldown co-immunoprecipitates GATA1/bio-GATA1 and TAF10 protein as detected by Western blot assays using M20 (GATA1) and 6TA 2B11 (TAF10) antibodies. BirA MEL cells were used as controls.

B. TAF10 (6TA 2B11) IP in bio-GATA1 O/E MEL cells co-immunoprecipitates GATA1 in the reverse IP.

Figure 9. ChIP assays of TAF10 and GATA1 in the *GATA1* locus in h^{FL} and h^{BL}

A TAF10 antibody (23TA 1H8 mAb) and a GATA1 antibody were used to immunoprecipitate the formaldehyde-crosslinked chromatin from h^{FL} and h^{PB} erythroid progenitor cells.

A. Primers for GATA1 binding sites of the human GATA1 locus, as indicated (palindromic GATA1 binding site and GATA1 binding site at -3kb region relative to the TSS), were used to estimate the relative fold enrichment (RFE) of TAF10 and GATA1 by qRT-PCR. A CD71 antibody (isotype control) was used for the mock IP and background enrichment was set to value 1.

B. Enrichment for TAF10 and GATA1 at both binding sites of interest is shown in independent experiments for fetal liver and adult blood.

Overview of all ChIP experiments. See panel A for positions of primers.

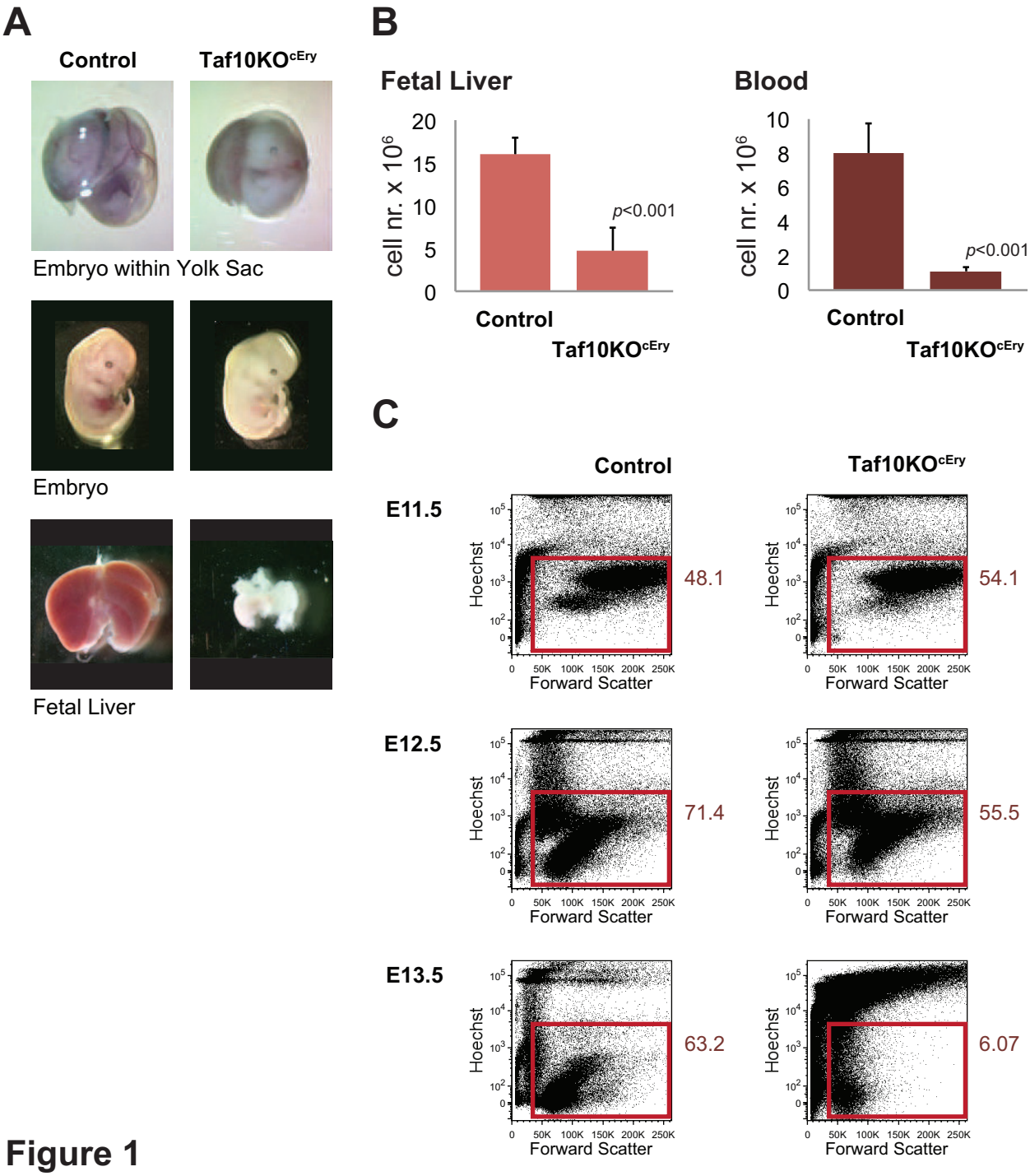


Figure 1

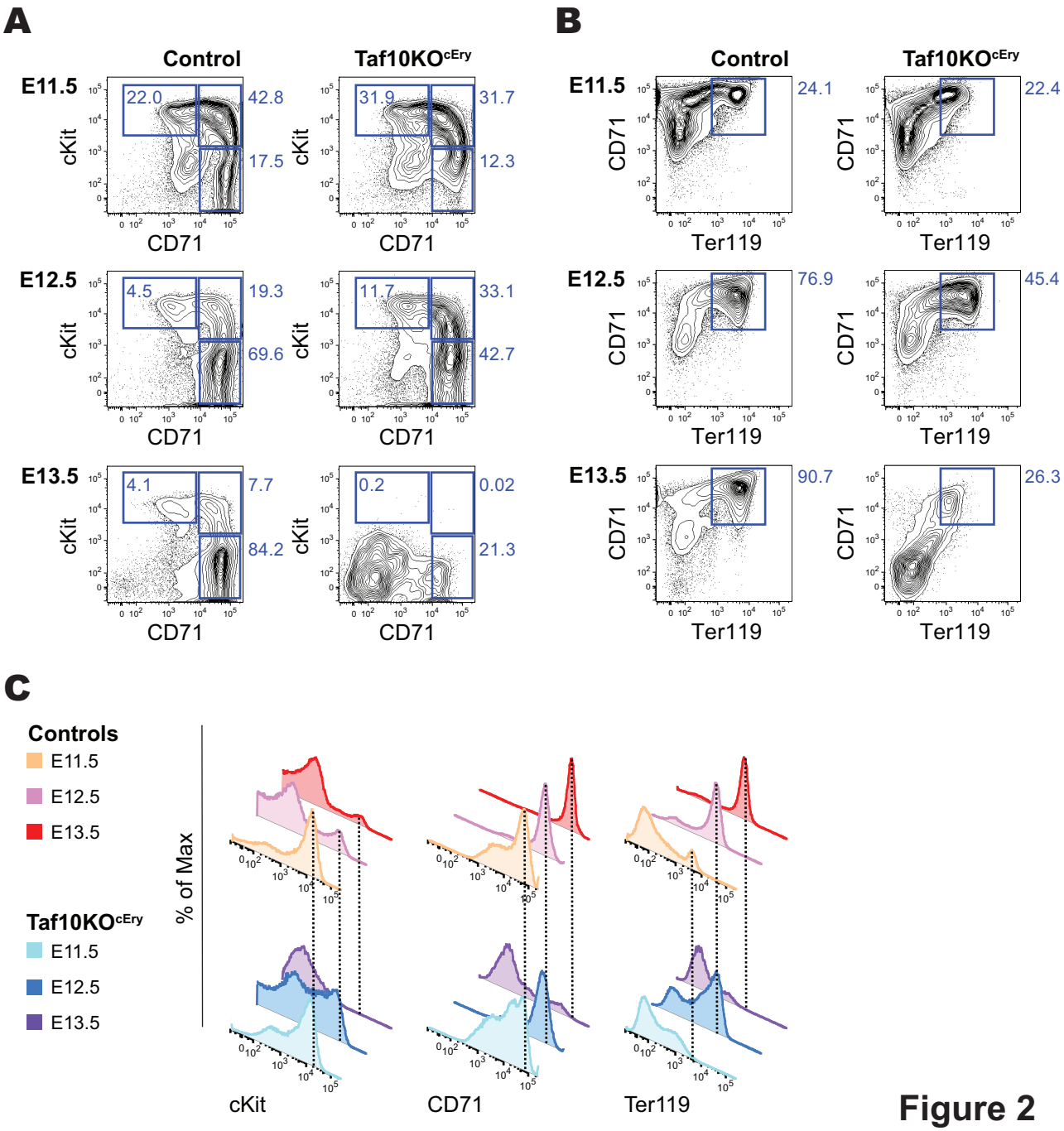


Figure 2

Molecular and Cellular Biology

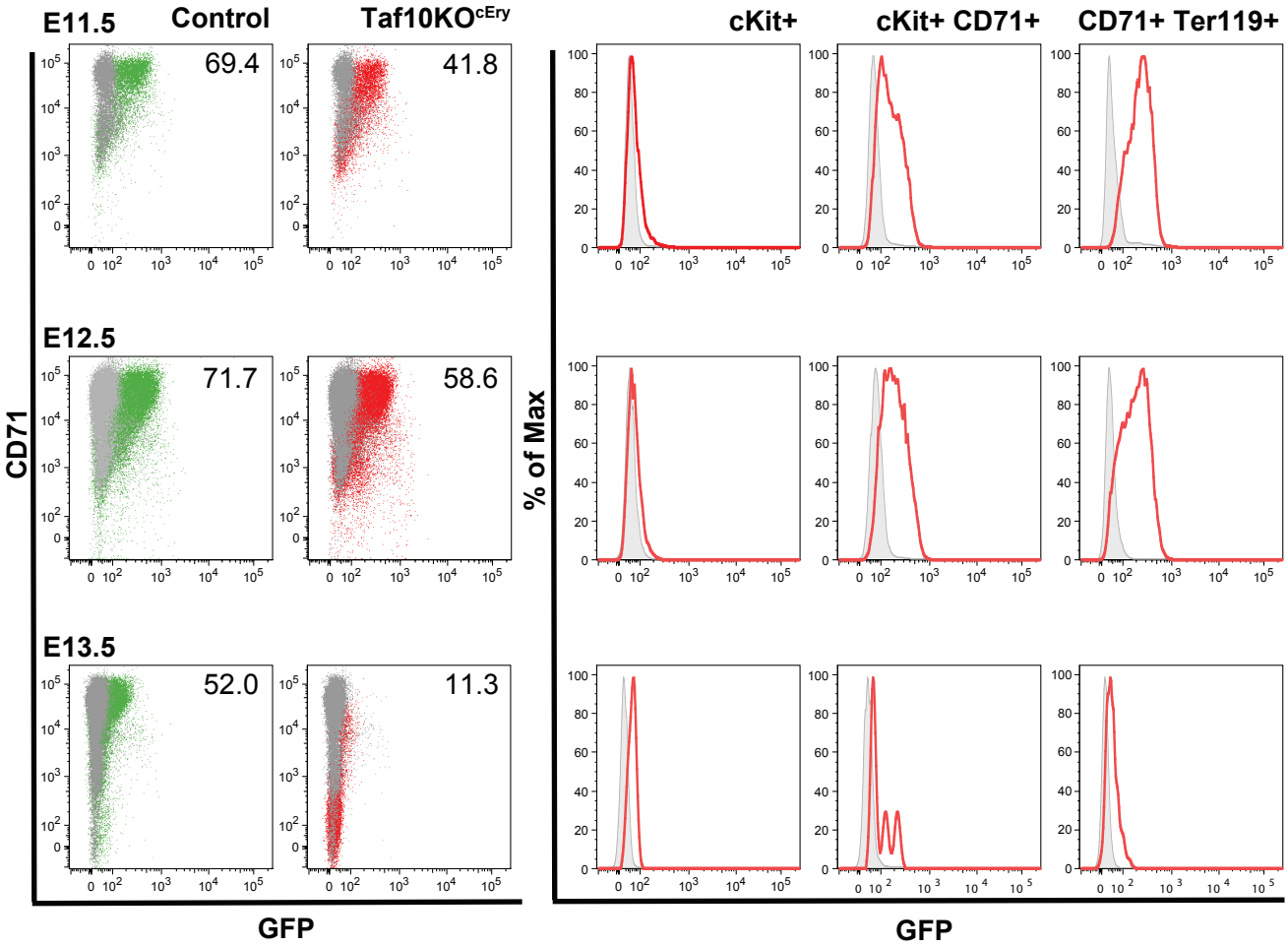


Figure 3

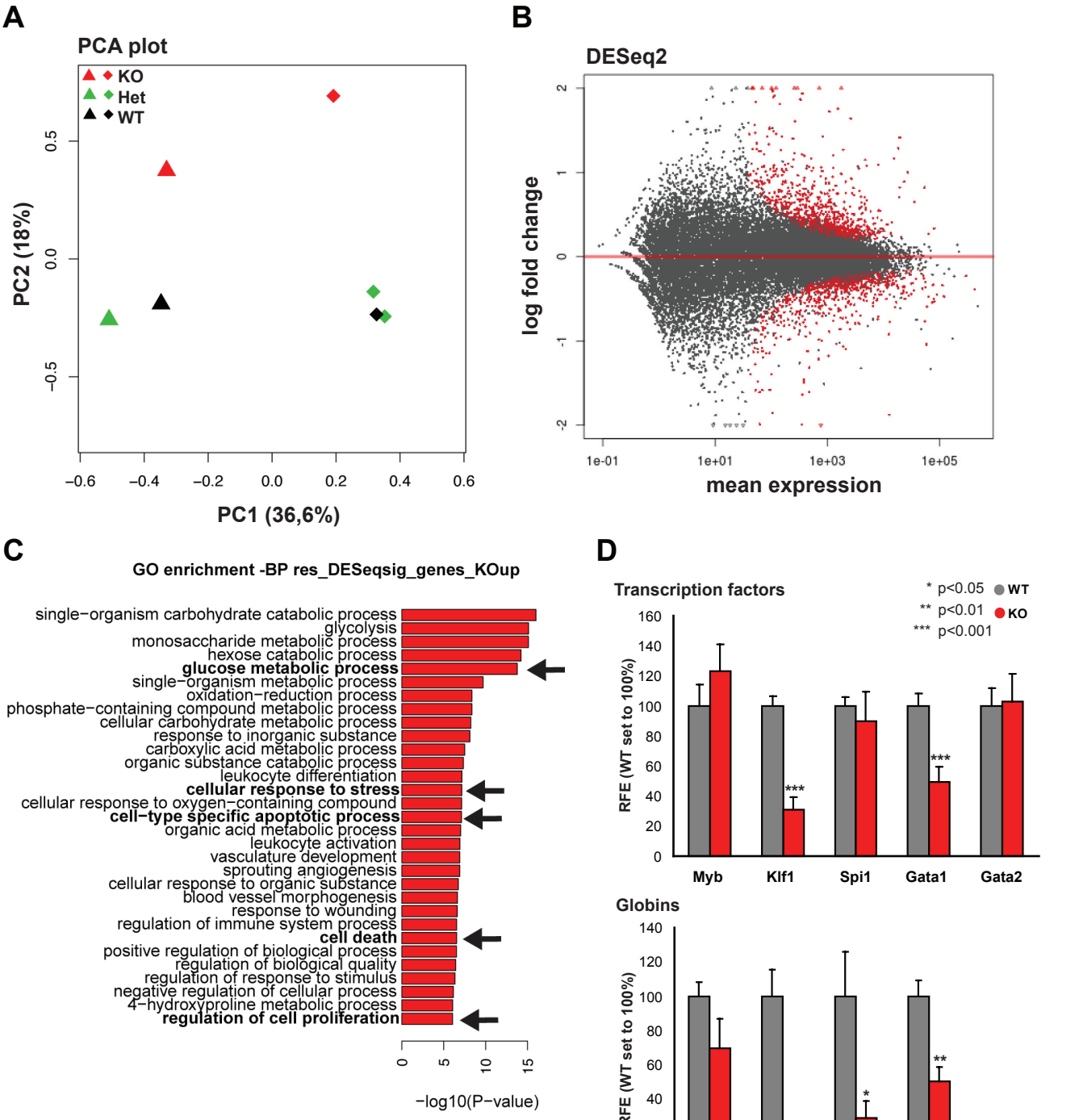
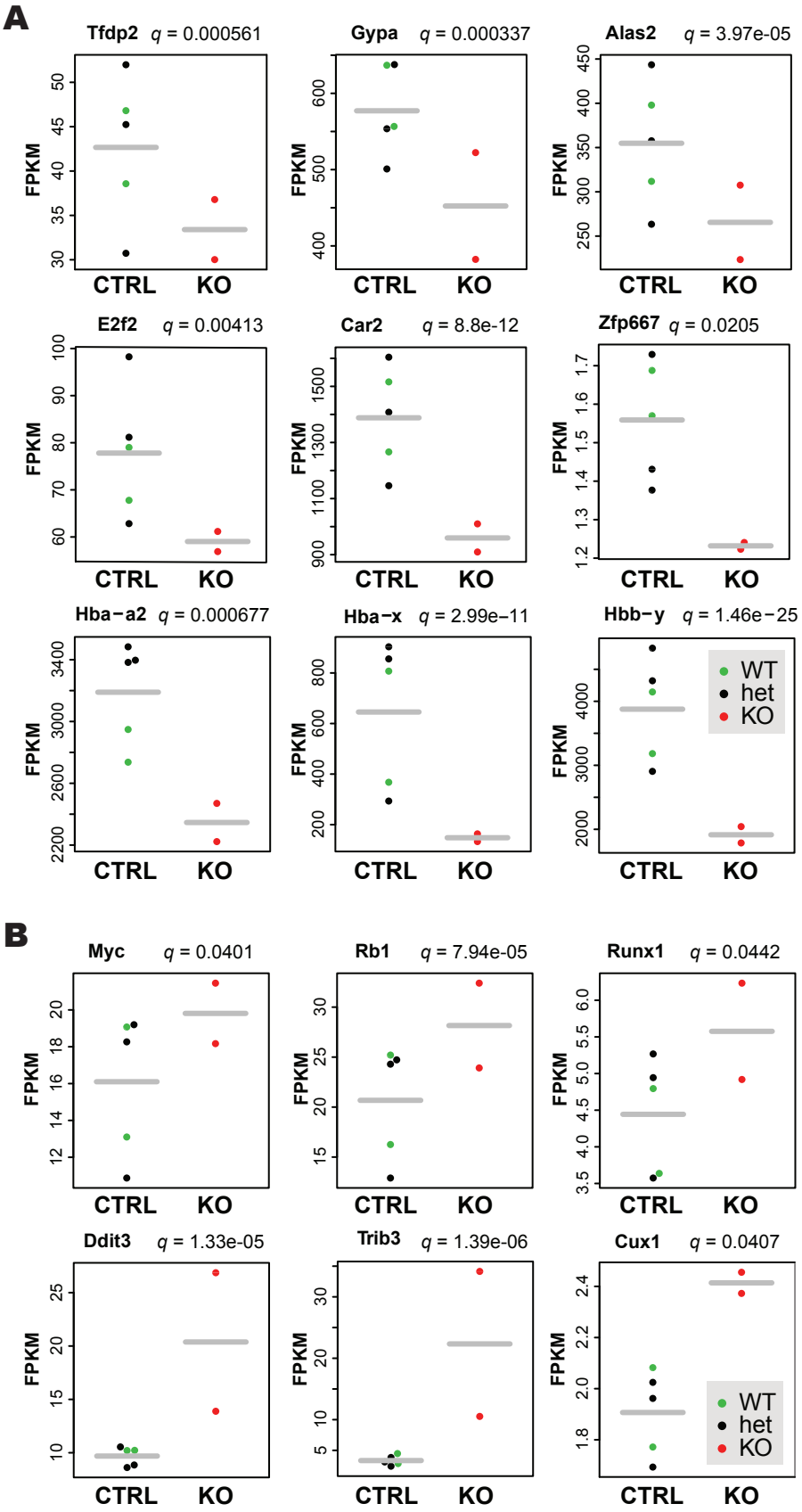


Figure 4



C

Downregulated Gata1 targets (q > 0,05)

Gene	Log2 Fold change	q-value
Gata1	-1,29E-01	2,92E-01
Nfe2	-2,26E-01	2,23E-01
Klf1	-2,69E-01	1,45E-01
Zbtb7a	-3,69E-01	1,02E-01
Hba-a1	-2,32E-01	6,59E-01
Hbb-bh1	-9,17E-01	7,47E-02
Hbb-bt	-1,33E+00	2,03E-01
Hbb-bs	-3,84E-01	7,37E-02

Upregulated Gata1 targets (q > 0,05)

Gene	Log2 Fold change	q-value
Gata2	2,93E-01	2,23E-01
Spi1	2,32E-01	2,90E-01

D

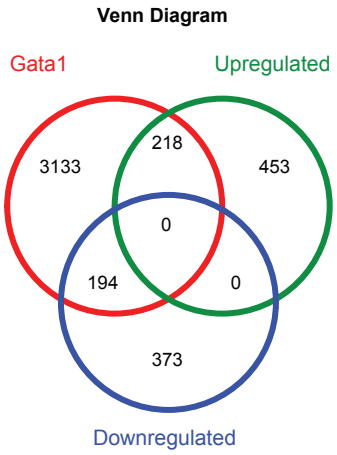


Figure 5

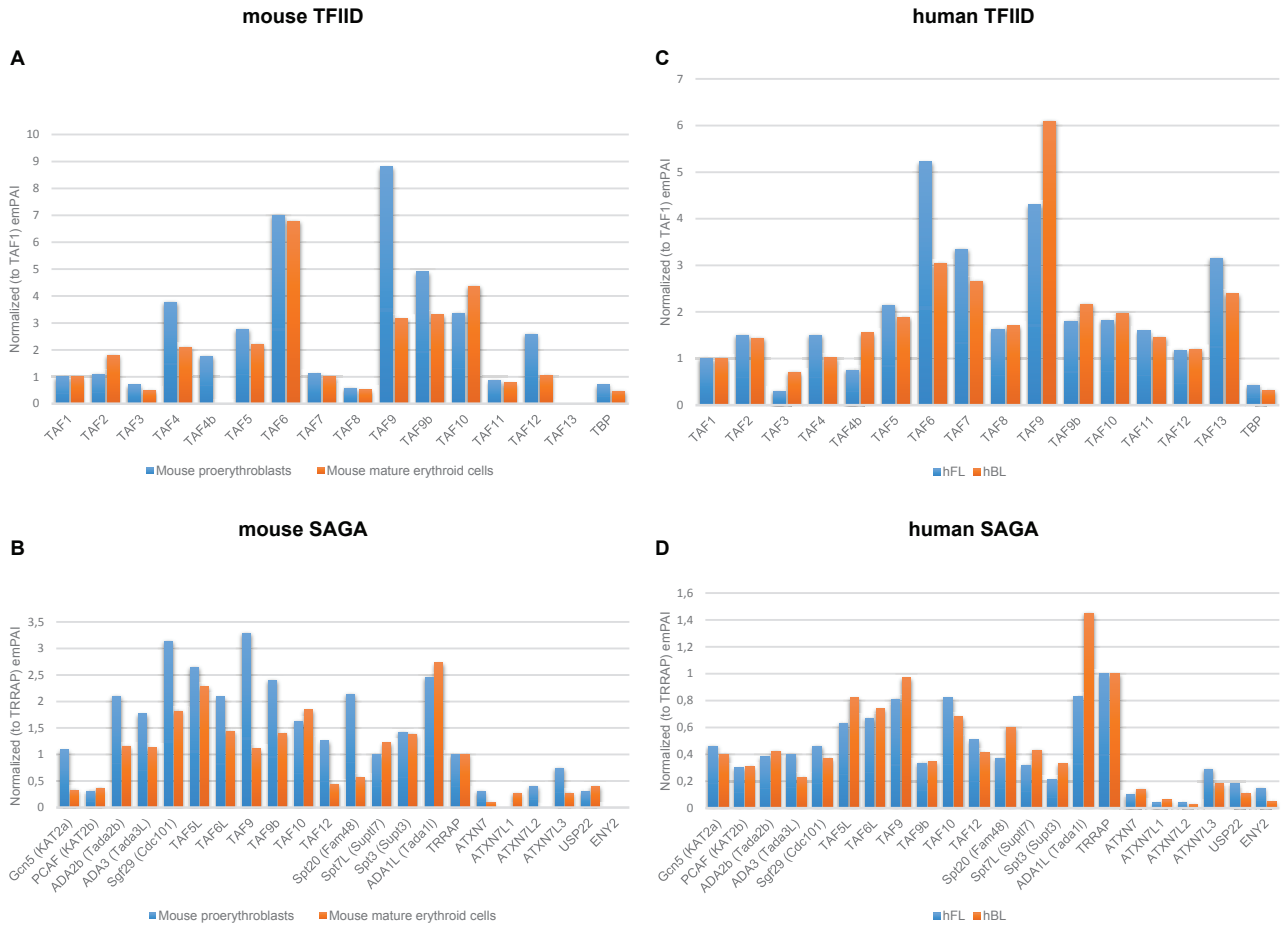


Figure 6

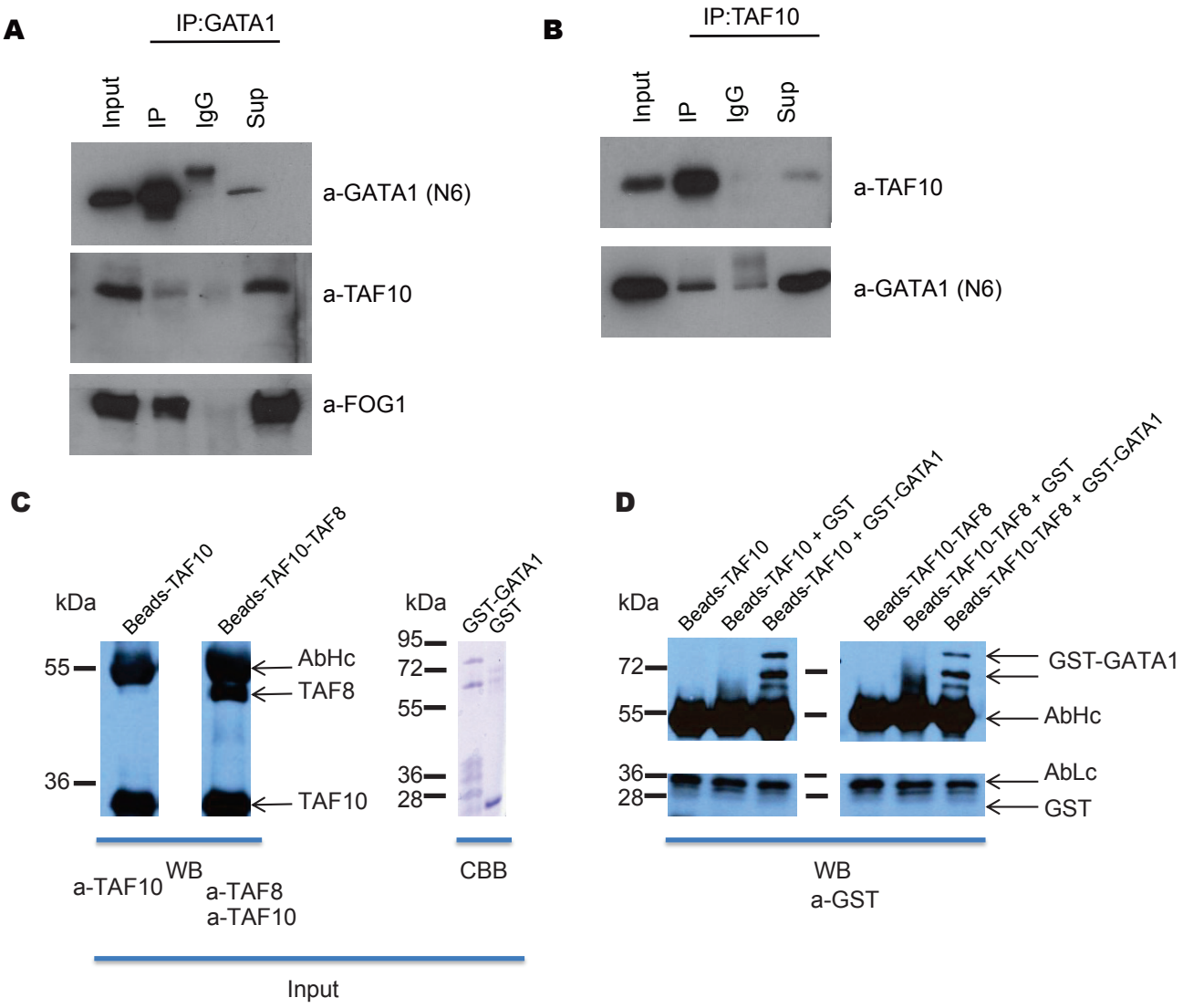
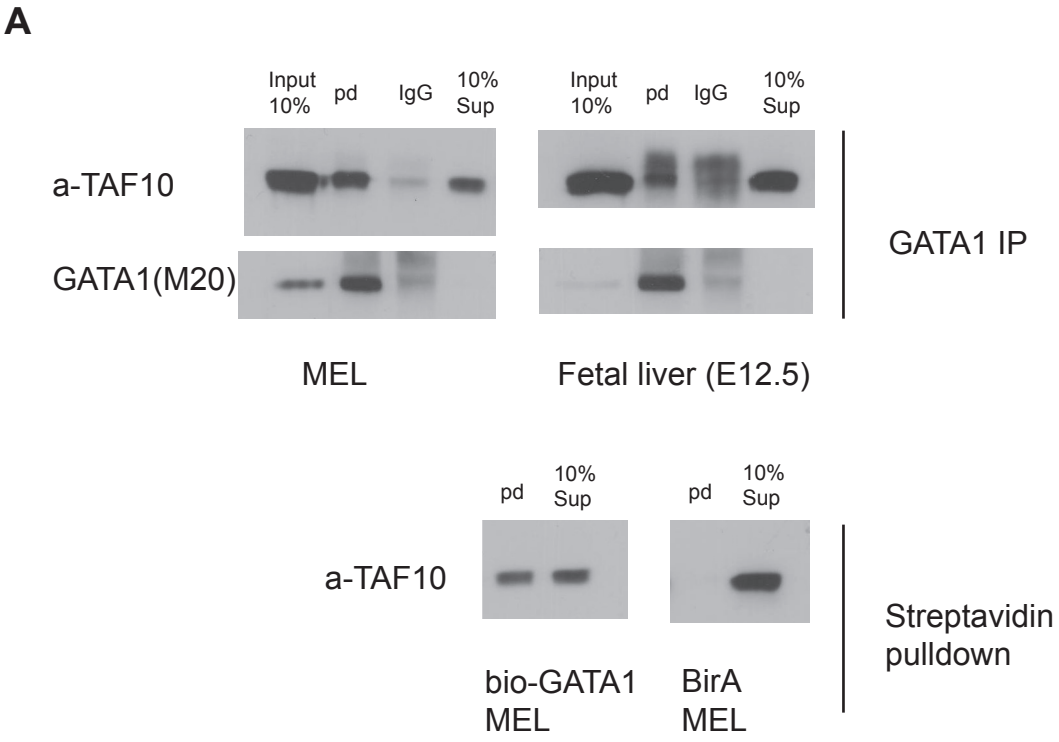


Figure 7



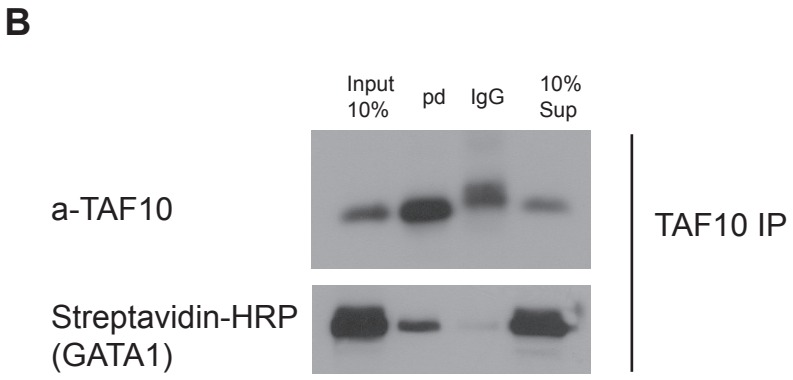


Figure 8

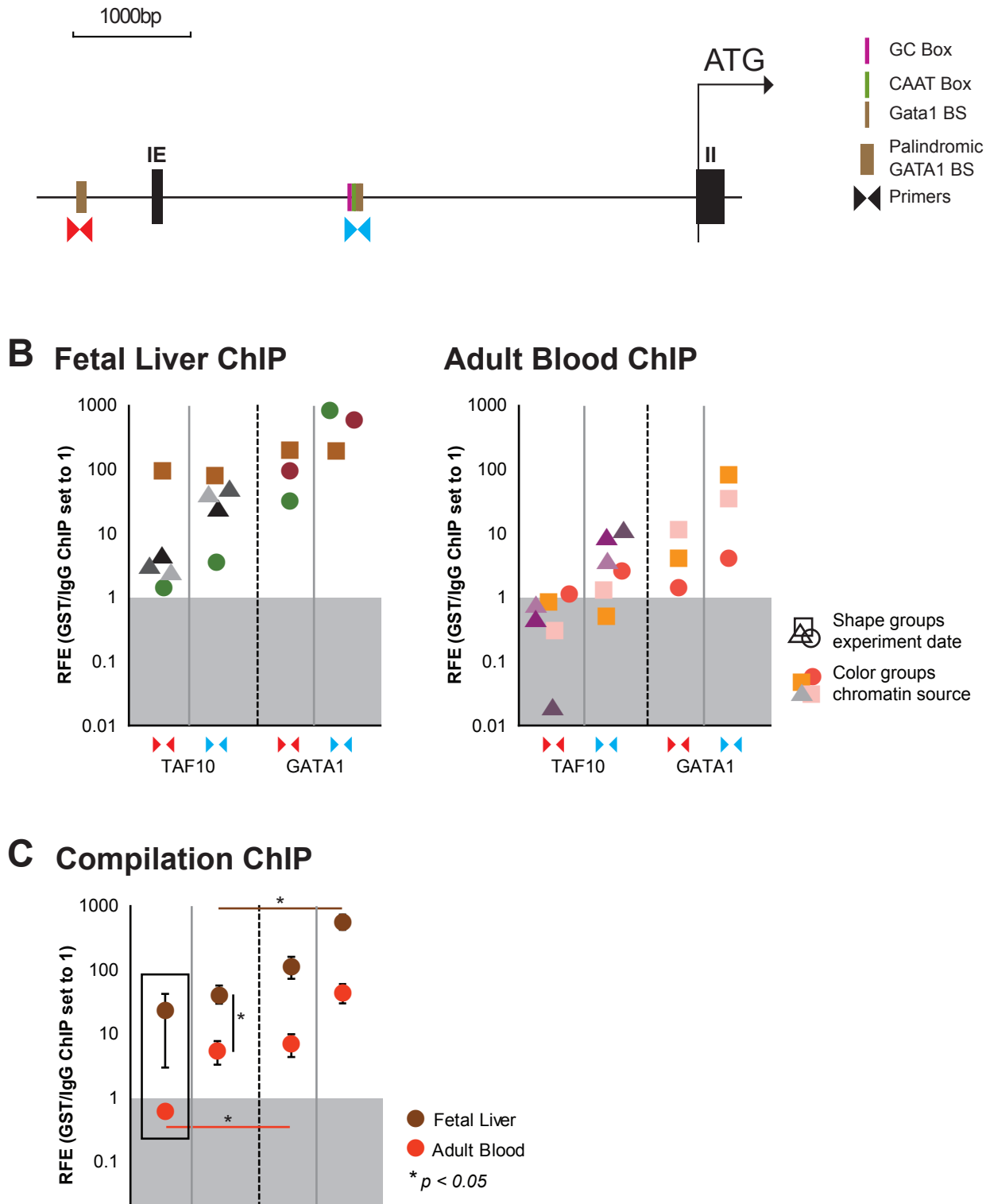


Figure 9

	Embryonic day	Control	Taf10KO^{cEry}	*p* value
% Live	E11.5	**39.55** ± *7.51*	**38.20** ± *22.49*	NS
	E12.5	**63.39** ± *5.94*	**52.53** ± *4.16*	< 0.005
	E13.5	**60.69** ± *3.21*	**6.55** ± *6.13*	< 0.001
% cKit	E11.5	**22.45** ± *5.99*	**35.15** ± *4.60*	NS
	E12.5	**5.61** ± *0.98*	**11.47** ± *1.24*	< 0.001
	E13.5	**3.85** ± *0.43*	**0.35** ± *0.40*	< 0.001
% cKit CD71	E11.5	**37.78** ± *6.96*	**20.95** ± *15.20*	NS
	E12.5	**16.30** ± *2.84*	**26.43** ± *5.67*	< 0.05
	E13.5	**8.03** ± *1.10*	**0.05** ± *0.03*	< 0.001
% CD71	E11.5	**17.10** ± *5.85*	**7.79** ± *6.38*	NS
	E12.5	**69.48** ± *3.19*	**50.95** ± *6.16*	< 0.01
	E13.5	**82.92** ± *2.09*	**14.50** ± *6.85*	< 0.001
% CD71 Ter119	E11.5	**29.00** ± *9.29*	**14.95** ± *10.54*	NS
	E12.5	**78.15** ± *1.99*	**55.93** ± *7.21*	< 0.01
	E13.5	**89.99** ± *1.64*	**14.57** ± *8.79*	< 0.001
cKit MFI	E11.5	**11121.63** ± *1768.39*	**10149.00** ± *2604.98*	NS
	E12.5	**2645.51** ± *126.30*	**3331.04** ± *618.45*	NS
	E13.5	**1440.50** ± *331.33*	**221.98** ± *188.65*	< 0.001
CD71 MFI	E11.5	**41095.75** ± *5407.63*	**19293.50** ± *12581.55*	NS
	E12.5	**36305.74** ± *2721.79*	**31207.59** ± *1445.95*	< 0.001
	E13.5	**43745.10** ± *2266.78*	**3034.50** ± *1274.59*	< 0.001
Ter119 MFI	E11.5	**1312.75** ± *250.17*	**647.50** ± *7.78*	< 0.001
	E12.5	**3175.30** ± *215.79*	**2194.38** ± *323.44*	< 0.005
	E13.5	**3734.50** ± *339.00*	**359.75** ± *92.00*	< 0.001

Table 1. Flow cytometry analysis of fetal livers
Statistical analysis derived from the flow cytometry analysis of fetal livers of Taf10KO^{cEry} and control embryos at E11.5, E12.5 and E13.5. At least 3 embryos are analyzed per genotype/stage.

Molecular and Cellular Biology

Protein mouse{m^{FLch}}/human{h^{FL},h^{PB}} /HELA TAF10IP	emPAI	Unique pept
TAF10	5.1/8.68(h^{FL}),4.29(h^{PB})/8.68	5/5(h^{FL}),5(h^{PB})/5
GATA-1$^{\#,\$}$	0.08/0.1(h^{FL})	1/1(h^{FL})
TAL-1$^{\#}$	0.1(h^{FL})	1(h^{FL})
LDB-1$^{\#}$	0.24(h^{FL})	3(h^{FL})
CNOT1	0.54/0.03(h^{PB})	31/2(h^{PB})
CNOT3	0.04	1
CNOT9	0.34/0.12	3/1
CNOT10	0.04/0.15(h^{FL})	1/1(h^{FL})
TRIM28	0.35/0.4(h^{FL}),0.44(h^{PB})/0.45	8/7(h^{FL}),8(h^{PB})/9
CBX3	2.01/0.88(h^{FL})	5/4(h^{FL})
CCAR1$^{\$}$	0.05/0.14(h^{FL})/0.26	2/5(h^{FL})/9
MED1$^{\$}$	0.04(h^{FL})/0.04	2(h^{FL})/3

Table 2. Selected activators and cofactors found in the TAF10IP by mass spectromerty (MS) Human (h^{FL}, h^{BL}) and mouse (m^{Flcl}: fetal liver cell lines) nuclear extracts of erythroid progenitor cells were used. Parameters of MS (emPAI and unique peptides) for each protein in different colors indicating species of origin and cell type are given. Symbols (#: proteins found in the same complex, $: reported to interact with each other).

TFIID[#] */SAGA*[#]	emPAI	Unique Pept
TAF1	0.02	4
TAF4a	0.1	2
TAF6	0.23	5
TAF5L	0.46	8
TAF9[#]	3.34	12
TAF9b[#]	0.94	6
TAF10[#]	0.17	1
TBP	0.35	4
PCAF(KAT2B)	0.12	5
ADA3(TADA3L)	0.25	5
Spt20(Fam48)	0.06	2
Spt3(Supt3)	0.19	3

Table 3. Mass spectrometry of GATA1 IP in MEL cells
TFIID and SAGA subunits were identified by MS after GATA1IP (bio-GATA1) on MEL nuclear extracts. Subunits that were found in BirA (Control) cells after streptavidin pulldown were subtracted from the list. Symbols (#: TFIID/SAGA shared subunits)